Leveraging Clinical Neuroimaging to Chart Brain Structure in 22q11.2 Deletion Syndrome

Author Byline:
Benjamin Jung, PhD[1,2,3]; J. Eric Schmitt, MD, PhD[2,4]; Jakob Seidlitz, PhD[1,2,3]; Jenna Schabdach, PhD[1,2,3]; Shivaram Karandikar, BA[1,2,3]; T. Blaine Crowley, BA[5]; Lena Dorfschmidt, PhD;[1,2,3], Ayan S. Mandal, MD, PhD[6,7]; Dabriel Zimmerman, BA[1,2,3]; Remo M.S. Williams, MS[1,2,3], Smrithi Prem, MD, PhD[1,2,3]; Elizabeth Levitis, PhD[1,2,3]; Margaret Gardner, BS[1,2,3]; Katherine Cyr, MS[1,2,3]; Viveknarayanan Padmanabhan, MS[8]; Jerome H. Taylor, MD[1,2,3]; Kosha Ruparel, MSE[2,3]; Rune Boen, PhD[9]; Carrie E. Bearden, PhD[9,10]; Christopher R. K. Ching, PhD[11]; Bogdan Pasaniuc, PhD[12,13]; Stewart Anderson, MD[1,2,3]; Daniel McGinn, MS[5]; Elaine Zackai, MD[5,12,13]; Beverly Emanuel, PhD[5,12,14]; Sarah Hopkins, MD, MSPH[5,15], Madeline Chadehumbe, MD[5,16]; Karen J. Low, PhD[17,18]; Tim J. Cole, PhD[19]; Richard A.I. Bethlehem, PhD[20]; R. Taki Shinohara, PhD[21,22]; J. William Gaynor, MD[23,24]; David R. Roalf, PhD[2,3], Raquel E. Gur, MD, PhD[1,2,3]; Donna M McDonald-McGinn, MS, CGC[5.12,14]; Aaron Alexander-Bloch, MD, PhD[1,2,3], on behalf of the ENIGMA 22q11.2 Deletion Syndrome Working Group

Affiliations:
1. Department of Child and Adolescent Psychiatry and Behavioral Sciences, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
2. Neurodevelopment & Psychosis Section, Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
3. Lifespan Brain Institute, Children's Hospital of Philadelphia and Penn Medicine, Philadelphia, Pennsylvania.
4. Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.
5. 22q and You Center, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
6. Department of Neurology, Mass General Brigham, Harvard Medical School, Boston, MA, United States
7. Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston, MA, United States
8. Translational Research Informatics Group, Department of Biomedical and Health Informatics | CHOP Research, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
9. Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, California
10. Department of Psychology, University of California, Los Angeles, Los Angeles, California
11. Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, California
12. Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
13. Department of Genetics, University of Pennsylvania, Philadelphia, Pennsylvania
14. Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
15. Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
16. Neurabilities Healthcare, Voorhees, New Jersey
17. Centre for Academic Child Health, Bristol Medical School, University of Bristol, Bristol, UK
18. Department of Clinical Genetics, UHBW NHS trust, Bristol, UK

19. UCL Great Ormond Street Institute of Child Health, London, UK
20. Department of Psychology, University of Cambridge, Cambridge, UK
21. Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
22. Center for Artificial Intelligence and Data Science for Integrated Diagnostics (Ai2D), Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
23. Division of Pediatric Cardiothoracic Surgery, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
24. Department of Surgery, University of Pennsylvania, Philadelphia, Pennsylvania

Corresponding Author:
Please address correspondence to Dr. Benjamin Jung.
Address: Richards Medical Research Building, 5th Floor, Suite b504, Philadelphia, PA, 19104
Email: jungb@chop.edu
Phone: 862-234-5864

Short Title: Clinical Imaging of Brain Structure in 22q11DS

Keywords: 22q11.2 deletion syndrome, normative modeling, copy number variation, electronic health records, schizophrenia, structural imaging

Abstract

Background: 22q11.2 deletion syndrome (22q11DS) is a common microdeletion linked to brain alterations and high psychiatric risk. However, prospective research studies often exclude severely affected individuals, potentially failing to capture the full neurobiological heterogeneity. Retrospective clinical MRI offers a powerful data source to address this gap.

Methods: This study utilized a cohort of 92 patients with 22q11DS (0.5-21 years, 49% female) and 252 matched controls. The authors applied normative modeling to calculate standardized deviation scores for imaging-derived phenotypes. They compared case-control effect size maps to data from the ENIGMA-22q consortium, correlated cortical deviation patterns with Allen Human Brain Atlas gene expression data, and generated 22q11DS-specific brain growth charts, mapping them to cognitive outcomes.

Results: 22q11DS was associated with widespread brain deviations, including smaller volumes (max Cohen's $d=-1.31$) and cortical surface area ($d=-0.71$) alongside increased cortical thickness ($d=0.39$). These findings were highly convergent with the ENIGMA-22q research cohort ($r=0.61$-0.87). A substantial portion of patients (40%) exhibited at least one global brain measure below the 2.5th percentile. Cortical deviations were spatially correlated with expression patterns of genes within the 22q11.2 locus. Deviations from the 22q11DS-specific growth charts were associated with clinical outcomes; smaller cerebellar volume relative to 22q11DS peers was associated with lower language scores ($p<0.03$).

Conclusions: This study validates the use of heterogeneous clinical imaging to robustly characterize brain structure in rare genetic disorders. Syndrome-specific growth charts provide a novel framework to quantify within-syndrome variability and demonstrate potential prognostic value by linking individual brain structure to cognitive outcomes.

Introduction

22q11.2 deletion syndrome (22q11DS; OMIM #188400, #192430) carries a profound burden of neuropsychiatric and cognitive symptoms (1). Notably, schizophrenia occurs in approximately 25% of adults with 22q11DS, alongside anxiety, attention-deficit/hyperactivity disorder (ADHD), and autism spectrum disorders (ASD) (1). The syndrome is also characterized by significant cognitive deficits spanning speech, language, executive function, and memory (2). These neuropsychiatric symptoms are co-occurrent with a host of multi-system symptoms including thymic dysplasia leading to immunodeficiency, congenital heart defects, and hypocalcemia (1).

The diverse clinical symptoms of 22q11DS are accompanied by widespread differences in brain structure compared with typically developing controls. Neuroimaging studies have consistently identified lower cortical, cerebellar, and subcortical gray matter volume (GMV), together with lower cortical surface area (SA) and higher cortical thickness (CT), in 22q11DS relative to controls (3–6). Furthermore, inter-individual differences in brain structure have been linked to both cognitive ability and risk for psychosis (4,6–8). Many of these structural differences were first robustly identified through prospective research studies and through mega-analyses conducted by the ENIGMA 22q11.2 Deletion Syndrome (ENIGMA-22q) working group (4,6,9).

Despite their advances, prospective research studies may not capture the full heterogeneity of 22q11DS. Research studies frequently employ exclusion criteria that remove individuals with severe neurodevelopmental deficits, such as impaired language, lower IQ scores, or an inability to tolerate research MRI scans. Because these features are common in 22q11DS, their exclusion leads to a substantial loss of phenotypic heterogeneity in study samples and to findings that may

not generalize to the entire 22q11DS population (1). For example, rates of incomplete hippocampal inversion were found to be higher in clinical samples of 22q11DS than research samples (10). Furthermore, existing neuroimaging studies have not typically included infants and toddlers, a critical gap given that neuroanatomical differences in 22q11DS are thought to emerge early in life (4–6).

Retrospective analysis of clinical imaging may offer a powerful complement to prospective research studies. Clinical scans may capture individuals often missed in traditional research: the more severely affected, the very young, and those who require sedation. Furthermore, recent advances in deep-learning-based image segmentation have reduced the disparity in quality between imaging-derived phenotypes (IDPs) obtained from clinical and research scans (11–14). The application of normative modeling to clinical MRIs further enables a critical analytic shift: moving beyond case-control comparisons to individual-level deviations from brain growth charts (15). By linking individual-level deviations to clinical outcomes, this framework provides a pathway towards imaging markers that capture the syndrome's profound clinical heterogeneity.

Here we demonstrate the utility of repurposing clinical neuroimaging in rare genetic disorder research. We assembled a cohort of 92 individuals with 22q11DS from the Children’s Hospital of Philadelphia (CHOP) and a reference sample of clinical controls with limited imaging pathology (16,17). By combining deep learning-based segmentation with normative modeling, our primary aims were to: (1) identify structural brain changes in 22q11DS clinical scans and compare them to ENIGMA-22q cohort; (2) quantify the degree of extreme deviation from brain charts of typical development; (3) identify transcriptomic correlates of these structural changes;

(4) generate 22q11DS-specific brain charts; and (5) identify associations between individual-level deviations and cognitive ability.

Methods and Materials

Study Cohorts and Design

Prospective data including the CHOP/Penn sample of ENIGMA-22q were collected with informed consent under protocols reviewed by CHOP and Penn Institutional Review Boards (IRB 10-007848 and IRB 812481 respectively). Retrospective use of clinical brain MRIs was reviewed by CHOP IRB 20–017870 and determined to be exempt from further oversight because it consisted of secondary use of existing data with low risk to human subjects and because informed consent could not realistically be obtained without biasing the sample.

The patient cohort was selected from individuals with 22q11.2 deletions and brain MRIs collected from 2000 - 2020. 22q11DS patients were further subsetted to include only those with deletions at positions LCR22A to LCR22D (i.e., A-D), which accounts for 85% of cases (1). After a multi-step process involving automated and manual quality control, the final sample consisted of 92 patients (**Supplemental Methods**, Figure S1). Clinical indications for scans included developmental delay, suspected seizure, and follow-up after cardiac surgery.

A reference group was assembled from a previously-described cohort of CHOP patients whose brain MRIs were determined to have limited or no imaging pathology (16,17). 252 clinical controls were matched to 22q11DS cases by sex, scanner, and closest age.

Primary findings were compared to the independent ENIGMA-22q dataset (242 individuals with A-D deletion, 277 typically developing controls) (see **Supplemental Methods**).

Image processing

De-identified scans were processed using the recon-all-clinical pipeline in FreeSurfer (v7.4.1), which enables cortical surface analyses across various scan resolutions, contrasts, and orientations (11–14). From these surface reconstructions, we extracted estimates for global and regional measures of GMV, sGMV, SA, and CT, as well as global white matter volume (WMV), ventricular volume, and cerebellum volume. Analyses of CT phenotypes were restricted to high-resolution magnetization-prepared rapid acquisition gradient-echo (MPRAGE) scans in patients aged two years or older due to suboptimal CT extraction outside of this group (N=40).

Severe brain pathologies are common in 22q11DS and may impact the ability to accurately segment the brain or introduce extreme IDPs that may bias findings towards structural pathologies as opposed to continuous brain variation. As such, each scan session from 22q11DS patients was evaluated by a board-certified neuroradiologist (J.E.S.) with expertise in 22q11DS to exclude patients with pathology that could interfere with surface reconstruction. Automated quality control measures from SynthSeg and surface Euler numbers (a proxy for scan quality) were further used to exclude low-quality segmentations (see **Supplemental Methods**) (12,18). Since many scan sessions contained multiple modalities, we derived a median IDP for each session across all scans that passed QC, rather than prioritizing a single scan, to maximize sample size and signal-to-noise, following a previously published protocol (12).

ENIGMA-22q research scans were processed using recon-all in FreeSurfer (v5.3) for consistency with prior studies (4,6) (see **Supplemental Methods**).

Normative Modeling

To quantify individual-level brain deviations, we constructed normative population models of brain development ("brain growth charts") using a large reference sample of 1,995 clinical scans (16,17). We used generalized additive models for location, scale, and shape (GAMLSS) to model growth of each IDP, with fixed effects to account for age and sex as well as a random intercept for scanner (19). Resulting models were applied to 22q11DS patients to generate standardized deviation scores (analogous to z-scores for normally distributed data). The Lifespan Brain Chart Consortium (LBCC) growth charts were used to derive ENIGMA-22q deviation scores (20). Normative modeling pipelines are discussed in detail in the **Supplement**.

Spatial Association with Gene Expression

To investigate potential molecular underpinnings of 22q11DS brain structure, we compared spatial patterns of regional deviation scores with brain-wide gene expression data from the Allen Human Brain Atlas (AHBA) (21). The primary principal component of gene expression in the brain was downloaded and parcellated using the *neuromaps* package in python (22). Vertex-based principal components were then aggregated using the Desikan-Killany atlas to generate gene expression maps at the same scale as brain deviation scores. Individual gene expression maps were generated and parcellated using the *abagen* package in python, with the primary analysis focusing on genes within the 22q11.2 deletion boundaries (23) (see **Supplemental Methods**).

Syndrome-Specific Growth Charts

To model brain growth specifically *within* the 22q11DS population, we applied the LMSz method (15). This approach uses standardized deviation scores (z scores) from growth charts of typical development as input and refines distributional parameters to create syndrome-specific growth charts.

Deviation scores from individuals with 22q11DS were used to estimate linear equations for the mean and scale of the 22q11DS brain growth charts (see **Supplemental Methods**). In the subset of patients with longitudinal imaging (N = 19), we included all available scans, with observations weighted by the inverse of the number of scans per participant to account for repeated measures (24). An iterative model selection procedure was used to exclude outliers (absolute residuals > 3.5) before reconstructing the final models. The final models yielded two outputs: (a) 22q11DS-specific deviation scores, quantifying how an individual deviates from their cohort norm, and (b) centile lines characterizing this distribution. By combining the SLIP and 22q11DS-adjusted models, we back-transformed the centile lines to generate population growth charts for 22q11DS.

Cognitive and Language Data

A subset of electronic health records (EHR) in the 22q11DS cohort contained assessments of intelligence quotient (IQ) and language ability. IQ and language scores in the EHR were matched to the closest clinical scan for evaluating inter-individual differences in brain structure and their relationship to cognitive outcomes.

IQ was assessed in 33 patients through the use of age-appropriate Weschler assessment scales, namely the Wechsler Preschool and Primary Scale of Intelligence (25,26), Wechsler Intelligence

Scale for Children (WISC) (25,27,28), Wechsler Abbreviated Scale of Intelligence (WASI) (29), and the Wechsler Adult Intelligence Scale (30,31). The Wechsler assessment scales evaluate intelligence quotients across multiple cognitive domains. Measures of IQ are further subdivided into subcategories of (VIQ) and (PIQ).

Language ability was evaluated through two independent scales: the Clinical Evaluation of Language Fundamentals (CELF; N = 48) (32,33) and the Preschool Language Scales (PLS; N = 58) (34–36). The CELF is a structured assessment administered to individuals between the ages of six and 21 that provides standardized scores for core language, receptive language, and expressive language. The PLS is a play-based assessment designed for individuals up to seven years old and provides standardized scores for language, auditory comprehension, and expressive communication.

Statistical Analysis

To assess case-control differences, each IDP was predicted using a linear fixed-effects model. The primary predictor was diagnosis, with covariates for age, sex, and the median Euler number (a proxy for scan quality). Age and sex were included to account for any residual biases not fully captured by normative modeling. From these models, we converted t-statistics to Cohen’s d effect sizes (37) and corrected all p-values for multiple comparisons using the Benjamini-Hochberg FDR procedure (38).

We used a permutation test to assess the statistical significance of the spatial correlation (Pearson's r) between the effect size maps from our clinical cohort and the ENIGMA-22q cohort. To do this, we first generated 1,000 null effect size maps using the Freedman-Lane procedure

(39). We then calculated correlations between these null effect size maps and the ENIGMA-22q map. The p-value was derived by comparing our observed correlation to this null distribution. As a confirmatory analysis, we also used a spatial "spin test," which produced convergent results (40).

We defined an "extreme" IDP as any value falling below the 2.5th or above the 97.5th percentile of the reference population. We then used a permutation test to determine if the rate of these extreme IDPs was significantly higher in patients with 22q11DS than would be expected by chance. This was done by creating a null distribution from 1,000 permutations of the case-control diagnosis labels and calculating the difference in extreme rates for each permutation.

To quantify the relationship between brain deviations and gene expression, we first calculated the Pearson correlation between the median brain deviation map and each gene's expression map. We performed this analysis separately for patients with 22q11DS and for controls. We then used a permutation test to determine if the brain-gene correlation was significantly different between the two groups. A null distribution was generated by calculating the difference in correlation coefficients across 1,000 permutations of the diagnosis labels.

In an exploratory analysis, we used linear regression to investigate the relationship between an individual's deviation from the 22q11DS-specific growth charts and their scores on IQ and language tests. Cortical thickness (CT) measures were excluded from these models due to the limited sample size for that measure.

Results

Cohort Characteristics

The primary clinical cohort consisted of 92 cases with 22q11DS (mean age=7.8 years; 49% female) and 252 matched clinical controls (Table 1). The ENIGMA-22q cohort (cases=242, controls=277) was significantly older (mean age=17.3 years) but did not differ significantly by sex or IQ (Table S1).

Global and Regional Deviations

Population modeling revealed widespread brain deviations in 22q11DS (Figure 1, Table S1). Patients with 22q11DS had smaller cerebellums (Cohen's d=-1.31, $p_{fdr}<10^{-25}$), GMV (d=-0.72, $p_{fdr}<10^{-9}$), sGMV (d=-0.63, $p_{fdr}<10^{-7}$), WMV (d=-0.72, $p_{fdr}<10^{-9}$), and SA (d=−0.71, $p_{fdr}<10^{-8}$) compared to controls. Conversely, patients had larger ventricular volumes (d=0.52, $p_{fdr}<10^{-6}$) and higher mean CT (d=0.39, p_{fdr}=0.02). Subcortical alterations were most prominent in ventral diencephalon and hippocampus (d=−0.75 to −0.83; $p_{fdr}<10^{-9}$) (Figure 1B). Cortical GMV and SA deviations were most pronounced in occipital lobe (peak d_{GMV}=−1.32, d_{SA}=−1.29; Figure 1C) with a gradient of attenuating effect sizes from caudal to rostral cortex. A prominent exception was higher insular volumes (d=0.28-0.39, p_{fdr}<0.02). In contrast, CT deviations were largest in medial orbitofrontal cortex (d=0.89-0.99, $p_{fdr}<10^{-7}$) and did not demonstrate a gradient. Focal decreases were identified in temporal and cingulate cortices (d=−0.38 to −0.89; p_{fdr}<0.05). Clinical MRI effect sizes showed a significant correlation with ENIGMA-22q effect sizes (Figure 1D, Figure S2, Table S2) assessed using both permutation testing and a cortical "spin" test (r=0.61–0.87; p_{perm} < 0.002, p_{spin} < 0.001) (40). While GMV and CT showed a strong 1-to-1 correspondence between cohorts, cortical SA effects were systematically more negative in ENIGMA-22q. Age- and sex-by-diagnosis interactions were largely non-significant (Figures S3-S4).

Extreme Phenotypes

A substantial number of individuals with 22q11DS exhibited IDPs at the extremes of the reference distribution (Table S4). We found that 40% of cases had at least one global IDP <2.5th percentile, while 21% had a global IDP >97.5th percentile. The most frequent finding was low cerebellum volume, observed in 28% of 22q11DS. These extreme deviations extended to the regional level: a significant proportion of 22q11DS had IDPs <2.5th percentile for the majority of GMV and SA IDPs, and CT IDPs >97.5th percentile (Figure 2). The regional distribution of these extreme deviations was moderately correlated with the ENIGMA-22q cohort for GMV ($r=0.66$, $p_{perm}=0.005$, $p_{spin}<0.001$) and SA ($r=0.81$, $p_{perm}=0.001$, $p_{spin}<0.001$) (Figures S5-6, Table S5).

Brain Deviation Patterns Correlate with Gene Expression

The median deviation maps of GMV and SA, but not CT, demonstrated significantly higher correlations with the first principal component of gene expression in 22q11DS compared to controls (Figure 3A; GMV: $r_{22q}=-0.50$, $r_{control}=0.15$, $p_{perm}=0.002$; SA: $r_{22q}=-0.48$, $r_{control}=0.19$, $p_{perm}=0.03$). We identified six genes within the 22q11.2 locus with expression correlated to median deviation scores (Figure 3B, Table S6): *TXNRD2*, *P2RX6*, *C22orf39*, and *DGCR8* with GMV and *TXNRD2*, *AIFM3*, and *ZNF74* with SA. All gene correlations with SA and the *TXNRD2* correlation with GMV were also found in the ENIGMA-22q cohort (Figure S7 and Table S7).

Characterizing 22q11DS-specific Brain Growth Charts

Syndrome-specific 22q11DS brain growth charts revealed developmental trends that were shifted compared to typical development (Figure 4A, Table S8). At the global level, models showed lower medians for GMV, sGMV, WMV, cerebellum volume and SA as well as higher medians for ventricular volume, but these shifts were stable across age and sex. At the regional level, only subtle age and sex variations were observed (Figure S8).

Association of Brain Deviations with Cognitive and Language Outcomes

In an exploratory analysis, an individual's deviation from 22q11DS brain charts was associated with verbal intelligence and language measures (Figure 4B). Higher verbal IQ (n=27) was nominally associated with intracranial volume (t=2.2, df=25, p=0.03), sGMV (t=2.2, df=25, p=0.04), and WMV (t=2.4, df=25, p=0.03). Higher proficiencies in both language assessments were associated with larger cerebellar volume (t=2.3, df=56, p=0.02 and t=2.6, df=46, p=0.01). Higher PLS scores were additionally associated with larger values across all global measures except ventricles (t=2.1-2.7, df=56, p<0.04).

Sensitivity Analyses

We performed several sensitivity analyses, which confirmed the robustness of reported findings (see **Supplemental Methods**). Sensitivity analyses included: controlling for intracranial volume; using only high-resolution MPRAGE scans; using alternative brain measures (i.e., centile scores and "raw" IDPs non-benchmarked to brain growth charts); and including patients with clinically significant pathology on neuroradiological review.

Discussion

In this retrospective study, we demonstrate that leveraging heterogeneous clinical MRI data for studying brain structure in rare genetic disorders is a powerful and scalable approach. By applying robust analysis pipelines to a large cohort with clinical MRIs, we charted brain deviations in 22q11DS starting in the first year of life, characterized by reduced SA and tissue volumes alongside increased CT. Our principal findings are fourfold: (1) the convergence of findings between clinical MRI and prospective research cohorts supports generalizability of research findings to a broader, more clinically representative 22q11DS population; (2) population modeling highlights both the magnitude and the heterogeneity of deviations from typical brain development quantified by reference brain charts; (3) brain deviations in 22q11DS mirror expression patterns in postmortem brain tissue of genes within the 22q11.2 locus; and (4) relative to 22q11DS-specific brain growth charts, individualized deviation scores may be associated with cognitive and language outcomes.

Critically, we highlight the wider potential for leveraging heterogeneous clinical MRI data to study rare genetic disorders. Given the large effect sizes typical of rare genetic disorders and availability of large samples of “clinical controls”, large samples of genetic carriers may not be required. For example, only five cases would be required to detect the effect size ($d=-1.31$) we report in cerebellar volume, with 80% power at $p<0.05$ (41). The convergence with prospective research studies (ENIGMA-22q) provides crucial support for the application of retrospective analysis of clinical MRIs to other neurogenetic variants, the vast majority of which lack the large research studies conducted in 22q11DS (42,43).

Even for the relatively well-characterized neuroanatomical signature of 22q11DS, namely globally reduced cortical SA and increased CT, the cellular underpinnings of macroscopic brain

alterations remain unknown. Reduced cortical SA may reflect impaired proliferation of neuronal progenitor cells during gestation, thereby limiting the overall size of the cortical sheet (6,44). Cortical thickening may be caused by over-proliferation of intermediate progenitor cells, atypical neuronal migration, or atypical cortical folding (44,45). Early developmental origins are well-founded in 22q11DS symptomatology, as dysmorphogenesis of the pharyngeal arch system during embryonic development has been implicated in 22q11DS cardiac and craniofacial defects (1).

Hypothesized early developmental origins are also supported by our results, which suggest that deviations from typical brain development occur early in life and remain relatively stable. We observed no significant age-by-diagnosis interactions, and 22q11DS growth charts were largely stable across age, even with the inclusion of children as young as six months. This provides cross-sectional evidence that the primary neuroanatomical signatures of 22q11DS are established early and proceed along a relatively parallel, albeit offset, trajectory. This is consistent with previous reports of SA, although some studies have reported age-dependent effects on CT which we did not detect, possibly due to our limited sample size for that measure (n=40) (6,45,46),

The spatial correspondence of 22q11.2 gene expression and brain deviations also supports a role for mitochondrial function in brain alterations. Our imaging-transcriptomic analysis implicates two genes involved in mitochondrial function: *TXNRD2* (essential for reactive oxygen species level regulation (47)) and *AIFM3* (involved in NADPH oxidation and mitochondrial apoptosis (48)). Of note, genes related to mitochondrial function were previously implicated by an imaging-transcriptomics study conducted in ENIGMA-22q (9). Moreover, diminished dosage of *Txnrd2* in wild-type mice has previously been shown to mirror neuronal and mitochondrial

deficits in layer 2/3 projection neurons of 22q11DS mouse models (47). Furthermore, introduction of *Txnrd2* rescues the neuronal antioxidant and dendritic growth deficit in vitro. While mitochondrial dysfunction has been linked to cognitive issues and impaired neurogenesis, the relationship between mitochondrial dysfunction and cortical surface area is unclear (49). Thus, more direct studies of mitochondrial function and brain structure in 22q11DS are warranted.

Our application of normative modeling demonstrates both the magnitude and the heterogeneity of deviations from reference brain charts. 40% of the 22q11DS cohort possessed at least one global brain measure in the bottom 2.5th percentile, highlighting the severe and pervasive neurobiological impact of the deletion. This level of extreme atypicality contrasts sharply with studies of idiopathic schizophrenia or ASD, which report overlapping extreme deviations in rarely more than 10% of individuals (50–52). However, even within this genetically homogeneous cohort, there was considerable heterogeneity. No individual brain region was considered extreme in a majority of patients, indicating that other factors—including epigenetics, environmental effects, and background polygenic variation—are likely significant contributors to brain structure variability in 22q11DS. Given that inter-individual differences in brain structure are associated with clinical outcomes (including language ability in this study and psychosis later in life (6)), continuing to unravel the multifactorial nature of brain heterogeneity in 22q11DS remains an important direction for future work.

A primary innovation of this work is the development of syndrome-specific brain growth charts, which support moving beyond group-level comparisons toward a more personalized clinical framework. Specifically, we identified a novel link between smaller cerebellar volume, relative

to 22q11DS brain charts, and lower language proficiency scores across two different assessments. This finding connects cerebellar hypoplasia, a common feature of 22q11DS (3), to speech and language deficits characteristic of the syndrome. With further validation, a clinician could potentially plot a child's cerebellar volume on such a chart; a value in a low percentile *for the syndrome* could serve as an empirical basis for placing patients into higher need treatment groups, ensuring more personalized treatment plans that optimize the use of scarce clinical resources. Future development could additionally link early imaging biomarkers to long term outcomes in 22q11DS. To this end, combining early-life clinical imaging with prospective research assessing clinical, social, and cognitive outcomes may highlight prognostic biomarkers.

Limitations

Our findings must be interpreted in the context of several limitations. While our analytical approach was designed to be robust to image heterogeneity, and the clinical MRI results were convergent with research data, we cannot rule out subtle effects due to greater image heterogeneity present in clinical scans. Furthermore, segmentation was not reliable for infants younger than 6 months, a critical neurodevelopmental period that remains a key target for future technical development. Additionally, our analyses are cross-sectional and cannot make causal inferences about developmental trajectories, for which longitudinal data are required. Cognitive assessments were not collected concurrently with imaging. Thus, associations between cognitive measures and brain structure may reflect long-term structural effects rather than capturing the relationship between brain structure and function at a given point during development. Finally, brain structure in 22q11DS may be mediated by co-occurrent symptoms or procedures. For example, although congenital heart defects in 22q11DS have not been associated with fetal

growth, corrective and palliative surgeries for congenital heart defects have been associated with impaired perioperative brain growth (53,54).

Conclusions

Notwithstanding these limitations, our study demonstrates a powerful and scalable approach to improve access to neuroimaging for studying rare genetic disorders. By combining large-scale clinical MRI data with advanced analytical techniques, we can overcome many traditional barriers to studying neurogenetic conditions. Syndrome-specific growth charts may be a critical next step toward developing targeted, early interventions to improve outcomes for individuals affected by 22q11DS and other rare genetic disorders.

Acknowledgments: This work was supported by National Institute of Mental Health grants R01MH133843 and R01MH134896. Please refer to the supplement for a list of collaborators from the ENIGMA 22q11.2 Deletion Syndrome Working Group.

Disclosures: AAB, RAIB and JS hold equity in and JS and RAIB are directors of Centile Bioscience. RTS has received consulting income from Octave Bioscience. All other authors have nothing to disclose.

References

1. McDonald-McGinn DM, Sullivan KE, Marino B, Philip N, Swillen A, Vorstman JAS, *et al.* (2015): 22q11.2 deletion syndrome. *Nat Rev Dis Primers* 1: 15071.
2. Gur RC, Bearden CE, Jacquemont S, Swillen A, van Amelsvoort T, van den Bree M, *et al.* (2025): Neurocognitive profiles of 22q11.2 and 16p11.2 deletions and duplications. *Mol Psychiatry* 30: 379–387.
3. Schmitt JE, DeBevits JJ, Roalf DR, Ruparel K, Gallagher RS, Gur RC, *et al.* (2023): A Comprehensive Analysis of Cerebellar Volumes in the 22Q11.2 Deletion Syndrome. *Biol Psychiatry Cogn Neurosci Neuroimaging* 8: 79–90.
4. Ching CRK, Gutman BA, Sun D, Villalon Reina J, Ragothaman A, Isaev D, *et al.* (2020): Mapping Subcortical Brain Alterations in 22q11.2 Deletion Syndrome: Effects of Deletion Size and Convergence With Idiopathic Neuropsychiatric Illness. *Am J Psychiatry* 177: 589–600.
5. Schmitt JE, Vandekar S, Yi J, Calkins ME, Ruparel K, Roalf DR, *et al.* (2015): Aberrant Cortical Morphometry in the 22q11.2 Deletion Syndrome. *Biol Psychiatry* 78: 135–43.
6. Sun D, Ching CRK, Lin A, Forsyth JK, Kushan L, Vajdi A, *et al.* (2020): Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Mol Psychiatry* 25: 1822–1834.
7. Gudbrandsen M, Daly E, Murphy CM, Blackmore CE, Rogdaki M, Mann C, *et al.* (2020): Brain morphometry in 22q11.2 deletion syndrome: an exploration of differences in cortical thickness, surface area, and their contribution to cortical volume. *Sci Rep* 10: 18845.

8. Schmitt JE, Yi JJ, Roalf DR, Loevner LA, Ruparel K, Whinna D, *et al.* (2014): Incidental radiologic findings in the 22q11.2 deletion syndrome. *AJNR Am J Neuroradiol* 35: 2186–91.

9. Forsyth JK, Mennigen E, Lin A, Sun D, Vajdi A, Kushan-Wells L, *et al.* (2021): Prioritizing Genetic Contributors to Cortical Alterations in 22q11.2 Deletion Syndrome Using Imaging Transcriptomics. *Cereb Cortex* 31: 3285–3298.

10. Roalf D, Atkins A, Czernuszenko A, Pecsok MK, McDonald-McGinn DM, Schmitt JE, *et al.* (2025): Presence, severity, and functional associations of incomplete hippocampal inversion in 22q11.2 deletion syndrome. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2025.04.009

11. Iglesias JE, Billot B, Balbastre Y, Magdamo C, Arnold SE, Das S, *et al.* (2023): SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry. *Sci Adv* 9: eadd3607.

12. Billot B, Magdamo C, Cheng Y, Arnold SE, Das S, Iglesias JE (2023): Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain MRI datasets. *Proc Natl Acad Sci U S A* 120: e2216399120.

13. Billot B, Greve DN, Puonti O, Thielscher A, Van Leemput K, Fischl B, *et al.* (2023): SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. *Med Image Anal* 86: 102789.

14. Gopinath K, Greve DN, Das S, Arnold. S, Magdamo. C, Iglesias JE (2023): Cortical analysis of heterogeneous clinical brain MRI scans for large-scale neuroimaging studies. *arXiv*.

15. Low KJ, Foreman J, Hobson RJ, Kwuo H, Martinez-Cayuelas E, Almoguera B, *et al.* (2024): The LMSz method - an automatable scalable approach to constructing gene-specific growth charts in rare disorders. *medRxiv* 2024.08.19.24312213.

16. Schabdach JM, Schmitt JE, Sotardi S, Vossough A, Andronikou S, Roberts TP, *et al.* (2023): Brain Growth Charts for Quantitative Analysis of Pediatric Clinical Brain MRI Scans with Limited Imaging Pathology. *Radiology* 309: e230096.

17. Zimmerman D, Mandal AS, Jung B, Buczek MJ, Schabdach JM, Karandikar S, *et al.* (2025): A systematic protocol to identify “clinical controls” for pediatric neuroimaging research from clinically acquired brain MRIs. *bioRxiv* 2025.06.25.661530.

18. Rosen AFG, Roalf DR, Ruparel K, Blake J, Seelaus K, Villa LP, *et al.* (2018): Quantitative assessment of structural image quality. *Neuroimage* 169: 407–418.

19. Rigby RA, Stasinopoulos MD, Heller GZ, De Bastiani F (2019): *Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R*, 1st ed. Chapman and Hall/CRC.

20. Bethlehem RAI, Seidlitz J, White SR, Vogel JW, Anderson KM, Adamson C, *et al.* (2022): Brain charts for the human lifespan. *Nature* 604: 525–533.

21. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, *et al.* (2012): An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 489: 391–399.

22. Markello RD, Hansen JY, Liu ZQ, Bazinet V, Shafiei G, Suarez LE, *et al.* (2022): neuromaps: structural and functional interpretation of brain maps. *Nat Methods* 19: 1472–1479.

23. Markello RD, Arnatkeviciute A, Poline J-B, Fulcher BD, Fornito A, Misic B (2021): Standardizing workflows in imaging transcriptomics with the abagen toolbox ((S. Jbabdi, T. R. Makin, S. Jbabdi, J. Burt, & M. J. Hawrylycz, editors)). *eLife* 10: e72129.

24. World Health Organization (n.d.): *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*. Retrieved August 25, 2025, from https://www.who.int/publications/i/item/924154693X

25. Wahlstrom D, Raiford SE, Breaux KC, Zhu J, Weiss LG (2018): The Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition, Wechsler Intelligence Scale for Children— Fifth Edition, and Wechsler Individual Achievement Test—Third Edition. *Contemporary Intellectual Assessment: Theories, Tests, and Issues, 4th Ed*. New York, NY, US: The Guilford Press, pp 245–282.

26. Wechsler D (2012, November 12): Wechsler Preschool and Primary Scale of Intelligence--Third Edition. American Psychological Association (APA). https://doi.org/10.1037/t15177-000

27. Wechsler D (2014): *Wechsler Intelligence Scale for Children*, 5th ed. Bloomington, MN: Pearson.

28. Wechsler D (2012, November 12): Wechsler Intelligence Scale for Children, Fourth Edition. American Psychological Association (APA). https://doi.org/10.1037/t15174-000

29. Wechsler D (2011): *WASI-II: Wechsler Abbreviated Scale of Intelligence*. PsychCorp.

30. Wechsler D (2019, February 11): Wechsler Adult Intelligence Scale--Third Edition. American Psychological Association (APA). https://doi.org/10.1037/t49755-000

31. Wechsler D (2012, November 12): Wechsler Adult Intelligence Scale--Fourth Edition. American Psychological Association (APA). https://doi.org/10.1037/t15169-000

32. Semel E, Wiig E, Secord W (n.d.): Clinical Evaluation of Language Fundamentals, Fourth Edition.

33. Semel E, Wiig E, Secord W (n.d.): Clinical Evaluation of Language Fundamentals, Fifth Edition.

34. Zimmerman IL, Steiner VG, Pond RE (n.d.): Preschool Language Scale, Third Edition. San Antonio, TX: Psychological Corporation.

35. Zimmerman IL, Steiner VG, Pond RE (n.d.): Preschool Language Scale, Fifth Edition. San Antonio, TX: Psychological Corporation.

36. Zimmerman IL, Steiner VG, Pond RE (n.d.): Preschool Language Scale, Fourth Edition. San Antonio, TX: Psychological Corporation.

37. Ben-Shachar MS, Makowski D, Lüdecke D, Patil I, Wiernik BM, Thériault R, *et al.* (2025, May 27): effectsize: Indices of Effect Size, version 1.0.1. Retrieved July 16, 2025, from https://cran.r-project.org/web/packages/effectsize/index.html

38. Benjamini Y, Hochberg Y (1995): Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57: 289–300.

39. Freedman D, Lane D (1983): A Nonstochastic Interpretation of Reported Significance Levels. *Journal of Business & Economic Statistics* 1: 292–298.

40. Alexander-Bloch A, Shou H, Liu S, Satterthwaite TD, Glahn DC, Shinohara RT, *et al.* (2018): On testing for spatial correspondence between maps of human brain structure and function. *Neuroimage* 178: 540–551.

41. Cohen J (2013): *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York: Routledge. https://doi.org/10.4324/9780203771587

42. Rees E, Kirov G (2021): Copy number variation and neuropsychiatric illness. *Current Opinion in Genetics & Development* 68: 57–63.

43. Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, *et al.* (2022): Rare coding variants in 10 genes confer substantial risk for schizophrenia. *Nature* 604: 509–516.

44. Rakic P (2009): Evolution of the neocortex: Perspective from developmental biology. *Nat Rev Neurosci* 10: 724–735.

45. Bagautdinova J, Zöller D, Schaer M, Padula MC, Mancini V, Schneider M, Eliez S (2021): Altered cortical thickness development in 22q11.2 deletion syndrome and association with psychotic symptoms. *Mol Psychiatry* 26: 7671–7678.

46. Jalbrzikowski M, Lin A, Vajdi A, Grigoryan V, Kushan L, Ching CRK, *et al.* (2022): Longitudinal trajectories of cortical development in 22q11.2 copy number variants and typically developing controls. *Mol Psychiatry* 27: 4181–4190.

47. Fernandez A, Meechan DW, Karpinski BA, Paronett EM, Bryan CA, Rutz HL, *et al*. (2019): Mitochondrial Dysfunction Leads to Cortical Under-Connectivity and Cognitive Impairment. *Neuron* 102: 1127-1142.e3.

48. Xie Q, Lin T, Zhang Y, Zheng J, Bonanno JA (2005): Molecular cloning and characterization of a human AIF-like gene with ability to induce apoptosis. *J Biol Chem* 280: 19673–19681.

49. Khacho M, Clark A, Svoboda DS, MacLaurin JG, Lagace DC, Park DS, Slack RS (2017): Mitochondrial dysfunction underlies cognitive defects as a result of neural stem cell depletion and impaired neurogenesis. *Hum Mol Genet* 26: 3327–3341.

50. Bedford SA, Lai MC, Lombardo MV, Chakrabarti B, Ruigrok A, Suckling J, *et al.* (2024): Brain-Charting Autism and Attention-Deficit/Hyperactivity Disorder Reveals Distinct and Overlapping Neurobiology. *Biol Psychiatry*. https://doi.org/10.1016/j.biopsych.2024.07.024

51. Haukvik UK, Wolfers T, Tesli N, Bell C, Hjell G, Fischer-Vieler T, *et al.* (2025): Individual-level deviations from normative brain morphology in violence, psychosis, and psychopathy. *Transl Psychiatry* 15: 118.

52. Wolfers T, Doan NT, Kaufmann T, Alnaes D, Moberget T, Agartz I, *et al.* (2018): Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. *JAMA Psychiatry* 75: 1146–1155.

53. Matthiesen NB, Agergaard P, Henriksen TB, Bach CC, Gaynor JW, Hjortdal V, Østergaard JR (2016): Congenital Heart Defects and Measures of Fetal Growth in Newborns with Down Syndrome or 22q11.2 Deletion Syndrome. *J Pediatr* 175: 116-122.e4.

54. Cromb D, Bonthrone AF, Maggioni A, Cawley P, Dimitrova R, Kelly CJ, *et al.* (2023): Individual Assessment of Perioperative Brain Growth Trajectories in Infants With Congenital Heart Disease: Correlation With Clinical and Surgical Risk Factors. *Journal of the American Heart Association* 12: e028565.

Tables and Figures

Table 1. Demographic characteristics

		22q11DS (N=92)	**CN** (N=252)	
Age (years)				
	Mean (SD)	7.8 (5.2)	8.8 (5.4)	t = -1.5, p = 0.1
	Median [Range]	7.3 [0.5, 20.3]	7.6 [0.8, 19.6]	
Sex				
	Female	45 (49%)	114 (45%)	$\chi^2 = 0.4$, p = 0.5
	Male	47 (51%)	138 (55%)	
Race				
	White	67 (73%)	143 (57%)	$\chi^2 = 8.1$, p = 0.04
	Black/AA	16 (17%)	64 (25%)	
	Asian	2 (2%)	5 (2%)	
	Other/Multiple	7 (8%)	40 (16%)	
Ethnicity				
	Non-Hispanic	81 (88%)	160 (63%)	$\chi^2 = 0.6$, p = 0.4
	Hispanic	11 (12%)	29 (12%)	
	Unknown	0 (0%)	63 (25%)	
IQ				
	Mean (SD)	75.0 (13.0)	-	-
	Median [Range]	74 [52, 103]	-	
	Missing	64 (70%)	-	

Data are presented as N (%) or mean (SD) and median [range]. P-values reflect the statistical comparison between the 22q11.2 deletion syndrome (22q11DS) patient group and the control group from the CHOP cohort. IQ scores were not available for the CHOP reference group. Abbreviations: CHOP, Children's Hospital of Philadelphia; CN, controls; IQ, intelligence quotient; SD, standard deviation.

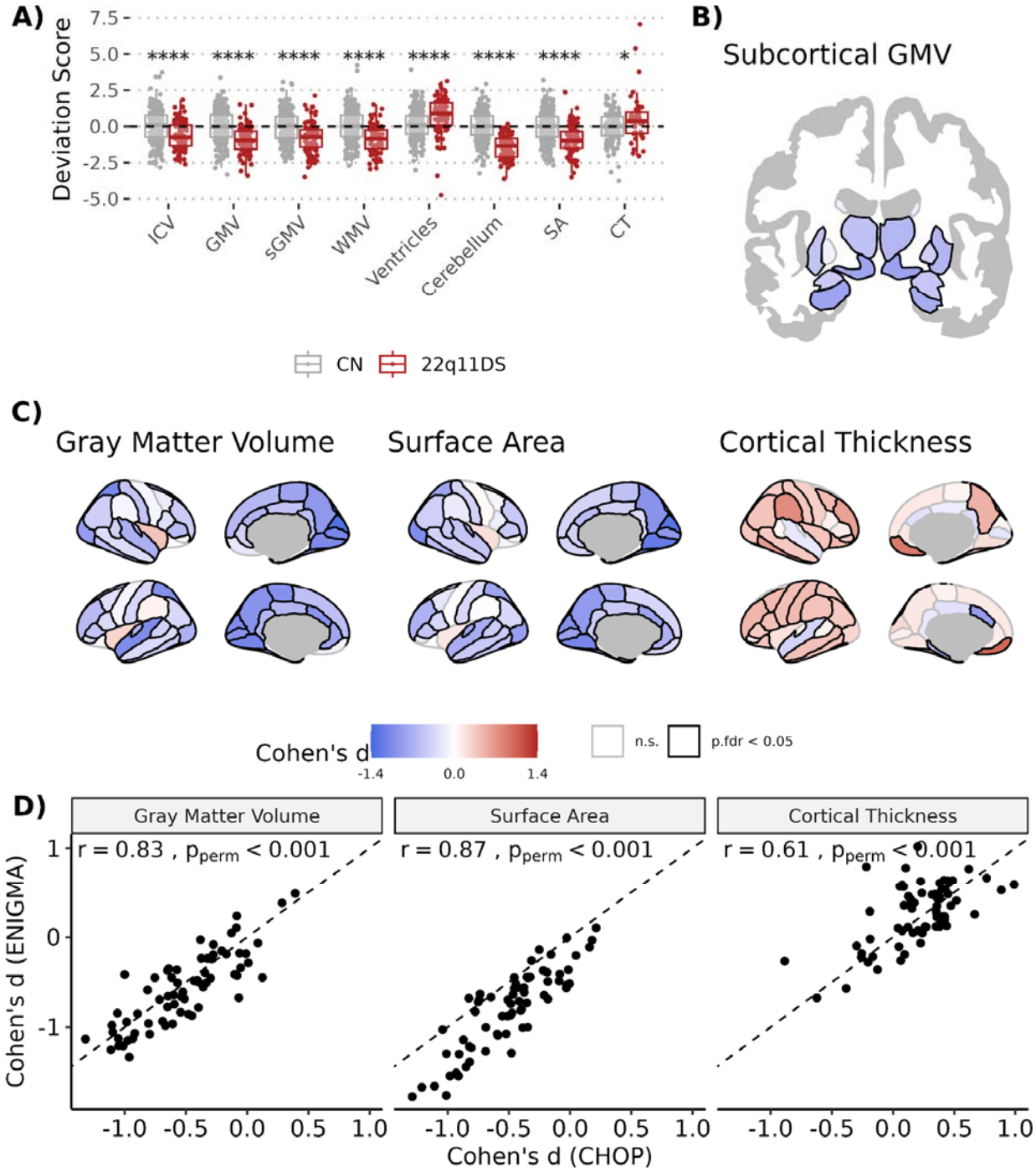

Figure 1. Effect sizes of 22q11.2 Deletion syndrome (22q11DS) on brain deviation scores. **(A)** Boxplots show individual standardized deviation scores for global brain measures in patients with 22q11DS (n=92) and controls (n=252). **(B** and **C)** Brain maps display Cohen's d effect sizes for regional deviations in **(B)** subcortical gray matter volume (GMV) and **(C)** cortical GMV, surface area (SA), and thickness (CT) imaging-derived phenotype (IDPs). Regions with negative deviation scores in 22q11DS (i.e., smaller IDPs in 22q11DS) are shown in blue, and regions with

positive scores are shown in red. Solid black outlines indicate regions with significant case-control differences. **(D)** Scatter plots show the correlation between Cohen's d effect sizes from the primary CHOP cohort (x-axis) and the independent ENIGMA-22q cohort (y-axis). All deviation scores were derived from population growth charts predicting IDPs based on age, sex, and scanner. P-values were corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) procedure. Asterisks indicate the level of statistical significance after correction. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * < 0.05. Abbreviations: IDP, imaging-derived phenotype, CT, cortical thickness; GMV, gray matter volume; ICV, intracranial volume; n.s., not significant; SA, total SA; sGMV, subcortical GMV; WMV, white matter volume.

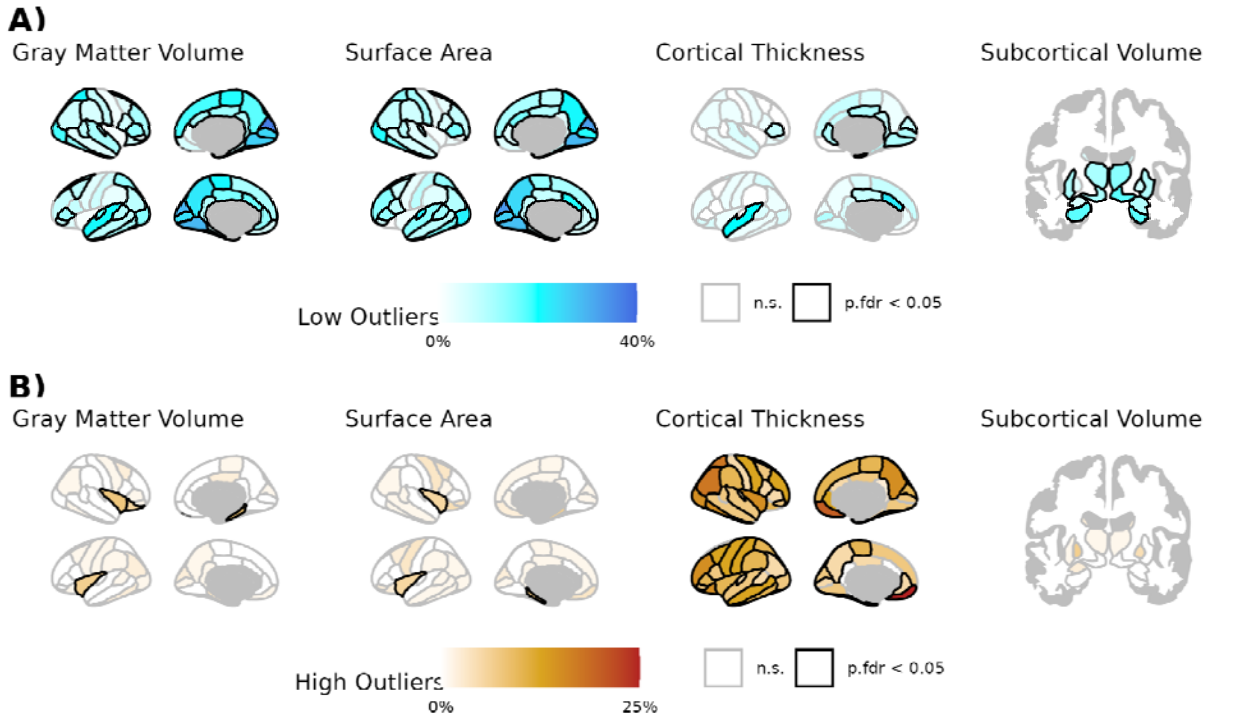

Figure 2. Extreme atypicality of imaging phenotypes is common in 22q11.2 Deletion syndrome (22q11DS). Brain maps showing the percentage of patients with atypically low scores for each imaging-derived phenotype (IDP) **(A)** and atypically high scores for each IDP **(B)**. Atypically low scores were defined as falling below the 2.5th percentile of the reference distribution, while atypically high scores were defined as above the 97.5th percentile. The statistical significance of regional enrichment for extreme phenotypes (indicated by solid black outlines) was assessed using a permutation test of case-control diagnoses (1,000 permutations). P-values were corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) procedure. n.s. indicates not significant.

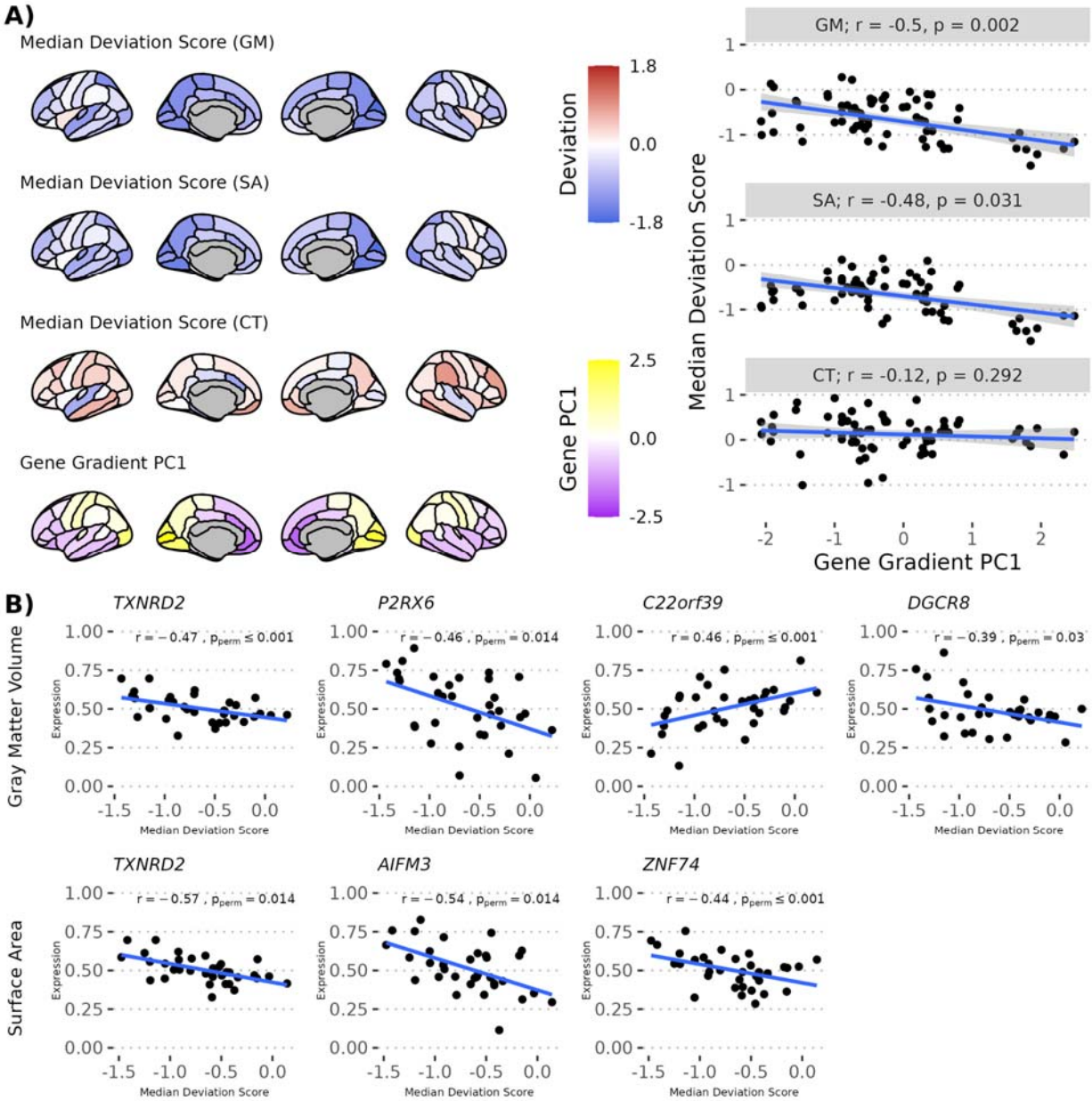

Figure 3. Spatial correlation between brain deviations in 22q11.2 Deletion syndrome (22q11DS) and expression of genes in the 22q11.2 locus in postmortem brain tissue (A) The spatial maps of median deviation scores for gray matter volume (GMV), surface area (SA), and cortical thickness (CT) were compared to the first principal component (PC1) of brain-wide gene expression from the Allen Human Brain Atlas. The scatter plots on the right show a significant negative correlation between the PC1 gradient and the deviation maps for both GMV ($r = -0.50$, $p_{perm} < 0.001$) and SA ($r = -0.48$, $p_{perm} = 0.03$) in 22q11DS. No significant correlation was observed for CT ($r = -0.12$, $p_{perm} = 0.3$). **(B)** Follow-up analyses show significant correlations between the GMV and SA deviation maps and the expression patterns of individual genes in the 22q11DS locus. Shown are the six 22q11DS genes that were significantly correlated with deviation maps after multiple comparisons correction. Significance was assessed by comparing

the difference in correlation coefficients of cases and controls to a null distribution across 1,000 permutations. Abbreviations: CT, cortical thickness; GMV, gray matter volume; PC1, first principal component; SA, surface area.

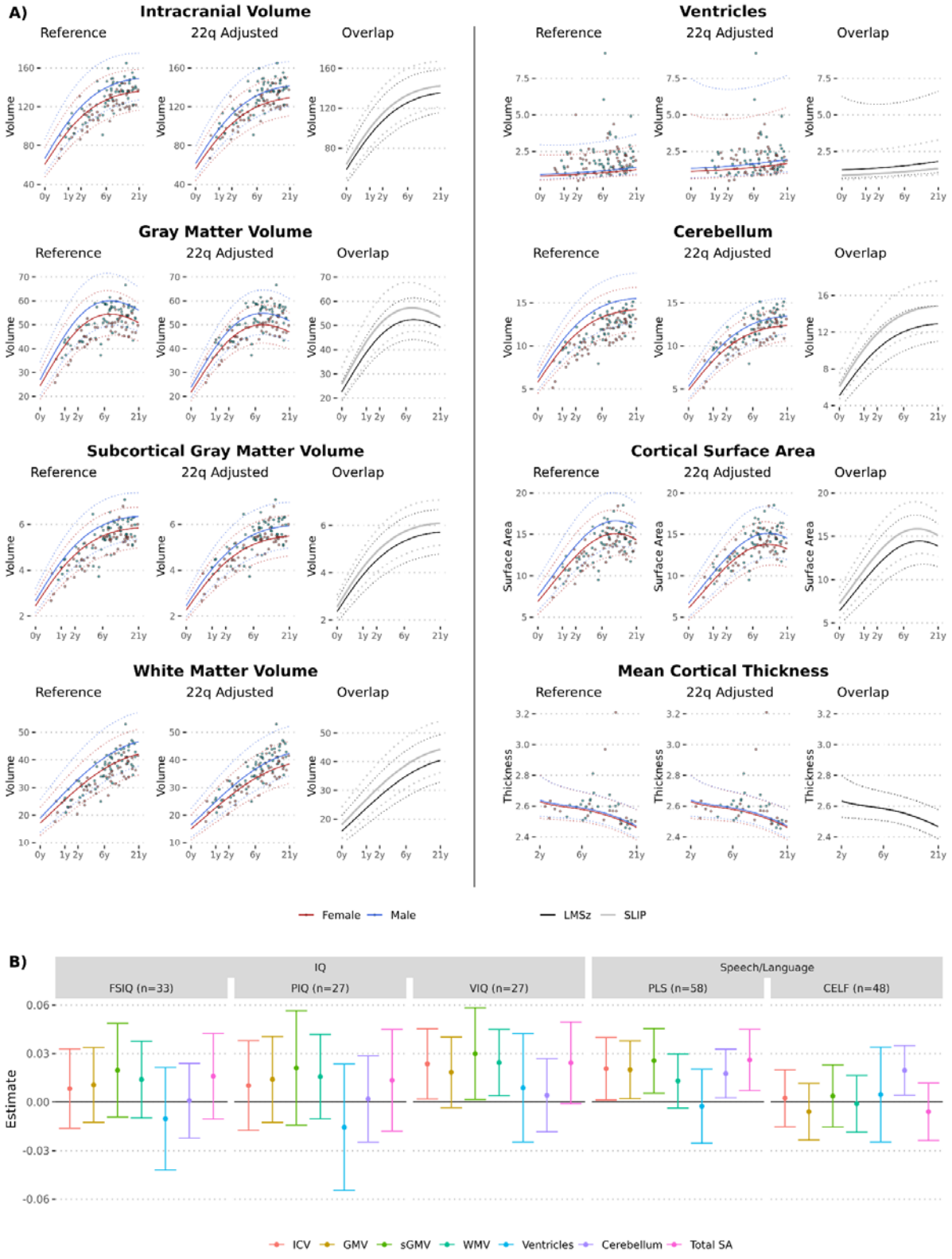

Figure 4. Syndrome-specific brain growth charts for 22q11.2 Deletion syndrome (22q11DS) and their association with cognitive outcomes. (A) Global brain growth charts for intracranial

volume (ICV), cortical gray matter volume (GMV), subcortical GMV (sGMV) and white matter volume (WMV). For each measure, the left panel ("Reference") shows individual data points from the 22q11DS cohort plotted against the reference population's growth curves. The middle panel ("22q Adjusted") displays the same data points against the newly generated growth charts that are specific to 22q11DS. The right panel ("Overlap") directly compares the median (solid lines) and 2.5th/97.5th centile (dashed lines) curves for both the reference (gray) and 22q11DS-specific models (black) averaged across sexes to visualize the shift in curves. (B) Exploratory analysis linking brain deviations to clinical outcomes. The points represent the estimated association (standardized beta coefficient) from linear models testing the relationship between an individual’s deviation from 22q11DS-specific growth charts and their scores on intelligence and language tests. Error bars indicate 95% confidence intervals. Cortical thickness measures were excluded from this analysis due to limited sample sizes. Abbreviations: CELF, Clinical Evaluation of Language Fundamentals; CT, cortical thickness; FSIQ, full-scale intelligence quotient; GMV, gray matter volume; ICV, intracranial volume; PIQ, performance intelligence quotient; PLS, Preschool Language Scales; SA, surface area; sGMV, subcortical GMV; VIQ, verbal intelligence quotient; WMV, white matter volume.