

Regional, circuit and network heterogeneity of brain abnormalities in psychiatric disorders

Received: 4 March 2022

Accepted: 13 July 2023

Published online: 14 August 2023

 Check for updates

Ashlea Segal ^{1,2}✉, Linden Parkes ^{3,4}, Kevin Aquino ^{1,2,5,6},
Seyed Mostafa Kia ^{7,8,9,10}, Thomas Wolfers ^{7,11,12}, Barbara Franke ^{13,14},
Martine Hoogman ^{13,14}, Christian F. Beckmann ^{8,15,16}, Lars T. Westlye ^{11,17,18},
Ole A. Andreassen ^{11,18}, Andrew Zalesky ^{19,20}, Ben J. Harrison ¹⁹,
Christopher G. Davey ²¹, Carles Soriano-Mas ^{22,23,24}, Narcís Cardoner ^{23,25,26},
Jeggan Tiego ^{1,2}, Murat Yücel ^{1,27}, Leah Braganza ¹, Chao Suo ^{1,2,28},
Michael Berk ^{29,30,31,32}, Sue Cotton ^{30,31}, Mark A. Bellgrove ¹,
Andre F. Marquand ^{7,8,15,33,34} & Alex Fornito ^{1,2,34}✉

The substantial individual heterogeneity that characterizes people with mental illness is often ignored by classical case–control research, which relies on group mean comparisons. Here we present a comprehensive, multiscale characterization of the heterogeneity of gray matter volume (GMV) differences in 1,294 cases diagnosed with one of six conditions (attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, depression, obsessive–compulsive disorder and schizophrenia) and 1,465 matched controls. Normative models indicated that person-specific deviations from population expectations for regional GMV were highly heterogeneous, affecting the same area in <7% of people with the same diagnosis. However, these deviations were embedded within common functional circuits and networks in up to 56% of cases. The salience–ventral attention system was implicated transdiagnostically, with other systems selectively involved in depression, bipolar disorder, schizophrenia and attention-deficit/hyperactivity disorder. Phenotypic differences between cases assigned the same diagnosis may thus arise from the heterogeneous localization of specific regional deviations, whereas phenotypic similarities may be attributable to the dysfunction of common functional circuits and networks.

The neurobiological mechanisms of mental illness are elusive. Thousands of neuroimaging studies have documented diverse brain changes associated with specific psychiatric diagnoses, and meta-analyses have identified the brain regions that are most consistently affected by each condition, revealing both diagnosis-specific and transdiagnostic effects^{1–7}. However, despite this substantial research effort,

pathophysiological processes are poorly understood and clinically useful biomarkers are lacking.

One reason for this limited progress may be a continued reliance on case–control designs, which compare group averages and ignore the considerable clinical heterogeneity often shown by individuals with the same diagnosis^{8,9}. Indeed, recent magnetic resonance imaging (MRI)

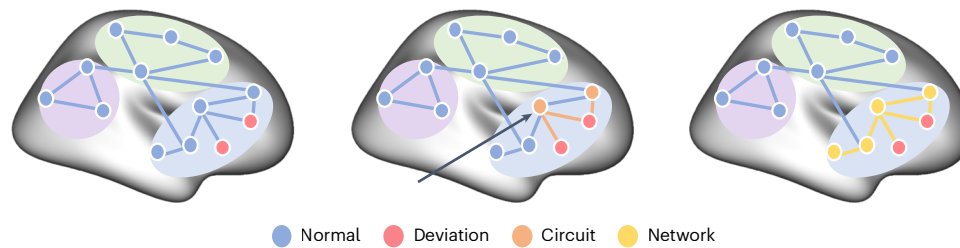


Fig. 1 | Characterizing neural heterogeneity at the level of brain regions, functional circuits and extended networks. A schematic showing how neural heterogeneity can be characterized at different scales. Nodes represent different brain regions, edges represent functional coupling (FC) between nodes and colored areas correspond to different functional networks of the brain. At the regional level (left), deviations from normative model predictions are localized to specific brain regions in each individual. Red nodes show the locations of such deviations mapped in two different people. A circuit-level analysis (middle)

reveals areas that are functionally coupled to the deviant loci. In this work, we define a functional circuit as the set of regions that show significant FC with a specific deviant region (orange). In this example, the two deviant areas are coupled to a common region (black arrow) despite being located in different areas themselves. These circuits can be embedded within extended networks (right) that include regions that may not be directly coupled to the deviant regions, but which nonetheless participate within the same functional system (yellow).

studies investigating person-specific patterns of brain deviations show that group average differences are not representative of individual cases¹⁰. These individual-specific inferences are often performed using normative modeling^{11,12}, which involves modeling normative expectations for a brain phenotype, such as gray matter volume (GMV), given an individual's age, sex or other relevant characteristics. The model predictions can then be used to define a normative range of variation against which new individuals are compared. Fitting the model to data for multiple brain regions yields a personalized deviation map quantifying the extent to which each person deviates from population norms, thus enabling identification of areas associated with unusually small or large phenotypic values, termed extreme deviations, in an individual. Normative modeling studies of diverse MRI-derived phenotypes in attention-deficit/hyperactivity disorder¹³ (ADHD), bipolar disorder (BP), schizophrenia^{10,14,15} (SCZ) and autism spectrum disorder^{16–18} (ASD) have found that, while cases often have more extreme deviations than controls (HCs; defined as the absence of any clinical diagnosis), the specific location of these deviations varies considerably across individuals with the same diagnosis.

This extreme regional heterogeneity of individual brain deviations aligns with the well-described clinical heterogeneity often associated with specific psychiatric diagnoses^{8,9}, but raises an important question: if cases show little overlap in the anatomical locations of their GMV deviations, what then explains phenotypic similarities between people assigned the same diagnostic label? It seems reasonable to assume that such similarities are driven by some common aspect of neural dysfunction across individuals, but the findings of normative modeling studies suggest otherwise.

One possible explanation is that these regionally heterogeneous deviations aggregate within common circuits or neural systems. The brain is a connected network, and pathological processes often affect distributed, interconnected systems^{19–21}, meaning that it is possible for deviations in disparate loci to impact the function of common, connected areas. This principle has been demonstrated by lesion network mapping studies of neurological patients sharing a common motor, perceptual or cognitive syndrome^{22,23}. Such patients typically show little overlap in the anatomical location of their lesions, but the lesioned sites are often functionally coupled to common areas (for a recent review, see ref. 24). The clinical expression of these syndromes is, therefore, more closely related to dysfunction of loci coupled to the damaged area, rather than dysfunction of the lesioned region itself.

In this Article, we considered whether a similar process is at play in psychiatric disorders by investigating whether anatomically heterogeneous regional brain deviations within psychiatric disorders are functionally coupled to common areas and networks.

We developed a new framework to integrate normative models of GMV with elements of lesion network mapping to map the functional circuits and extended networks within which regional GMV deviations are embedded. We used this approach to derive a multiscale characterization of neural heterogeneity across 1,294 individuals diagnosed with one of six disorders: ADHD, ASD, BP, depression (MDD), obsessive–compulsive disorder (OCD) and SCZ. Inspired by studies of patients with brain lesions^{22–24}, we tested the hypothesis that anatomically heterogeneous regional GMV deviations in each disorder are functionally coupled with common sites, either within a functional circuit or an extended functional network (for a schematic explanation, see Fig. 1). Our transdiagnostic, multiscale approach allowed us to comprehensively understand the extent of neural heterogeneity within each disorder while also revealing commonalities and differences between disorders.

Results

Sample characteristics

We examined data for 1,465 HCs (54.47% male) and 1,294 cases, taken from 14 different studies and 25 different scan sites. The clinical sample comprised 202 individuals with ASD (100% male), 153 individuals with ADHD (41.18% male), 228 individuals with BP (47.37% male), 161 individuals with MDD (34.16% male), 167 individuals with OCD (50.30% male) and 383 individuals with SCZ (62.14% male). The scanner details, sample size and demographic characteristics of each scan site, after various exclusions based on data quality and other criteria (Methods), are presented in Supplementary Table 1 (for age distributions, see Supplementary Fig. 1).

Normative modeling

We used an established pipeline to obtain voxel-wise estimates of GMV in each participant^{25,26} (Methods), which we aggregated into regional estimates for 1,032 brain areas^{27,28} (1,000 cortical and 32 subcortical regions; Fig. 2a). We then trained a normative model based on hierarchical Bayesian regression²⁹ (HBR), fitted separately in each region, in a training set comprising 1,196 HCs (55.02% male; Fig. 2a; HC_{train}) to establish a normative GMV range given an individual's age, sex and scan site (Fig. 2b; for model fit statistics and site effects, see Supplementary Fig. 3 and Supplementary Table 2). HC_{train} spanned the age range of cases, enabling predictions for people aged between 18 and 64 years (Supplementary Fig. 2). The remaining 269 controls (52.04% male; age range 18–62 years) were held out as a test set (HC_{test}) to establish a normative benchmark for comparison with each clinical group (Fig. 2c). For each individual in the clinical groups and HC_{test}, we quantified the degree to which regional GMV estimates deviated from normative model predictions as a z score (termed Deviation z maps;

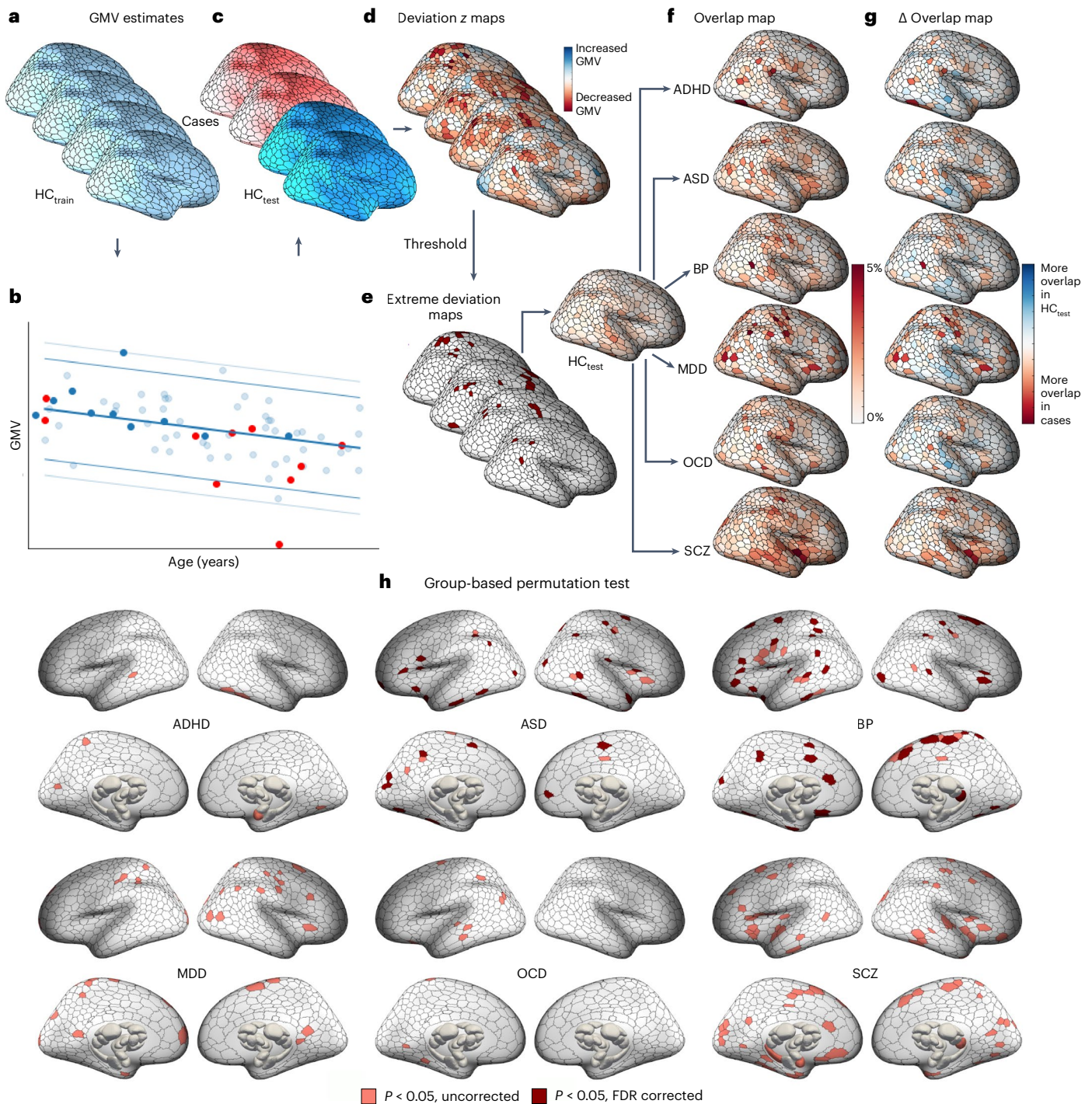


Fig. 2 | Regional heterogeneity of extreme negative GMV deviations in each disorder. **a–f**, Workflow for characterizing regional-level heterogeneity. GMV maps for each individual were parcellated into 1,000 cortical and 32 subcortical regions (**a**). The training dataset, HC_{train} , was used to train a normative model to make predictions about regional GMV values given an individual's age, sex and scan site (**b**). The predictions for held-out controls (HC_{test}) and cases were then compared with empirical GMV estimates. Model predictions for one region, showing individuals in the training set (HC_{train} ; light blue) and the held-out control (HC_{test} ; dark blue) and clinical groups (red). Solid and dashed lines indicate the 99th and 95th centiles, respectively (**c**). For each individual, deviations from model predictions were quantified as a deviation z map (**d**). This deviation map was then thresholded at

$z > |2.6|$ to identify extreme deviations (**e**). For the HC_{test} and each clinical group, we quantified the proportion of individuals showing an extreme deviation in a given brain region, yielding an extreme deviation overlap map (**f**). We subtracted the HC_{test} overlap map from each clinical group's overlap map to obtain an overlap difference map (Δ overlap map) for each clinical group and then evaluated the magnitude of this difference (for details, see Extended Data Fig. 3a–d) (**g**). Cortical and subcortical surface renderings showing regions with significantly greater overlap of extreme negative GMV deviations in cases compared with controls, as identified using group-based permutation tests (pink corresponds to $P_{uncorrected} < 0.05$, red corresponds to $P_{FDR} < 0.05$; two tailed, cases > controls) (**h**). Data used to generate this figure can be found in Supplementary Data 1 (Regional_neg_thr26).

Methods and Fig. 2d), with extreme deviations defined as $z > |2.6|$ (Fig. 2e), corresponding to $P < 0.005$, uncorrected, as per prior work¹⁰ (for a discussion of thresholding, see Methods).

Supplementary Table 3 shows that, across all groups, >75% of participants show at least one extreme negative deviation. The proportion of such participants was ~5–12% higher in MDD, SCZ, BP and OCD,

compared with controls. Over 65% of participants showed at least one positive extreme deviation, and the proportion of such participants was -4–10% higher in ADHD, ASD, BP and OCD compared with controls. People with BP, MDD, OCD and SCZ showed a higher burden of extreme negative deviations, defined as the total number of extreme deviations identified in each person, compared with HC_{test} ($P = 0.003$; Extended Data Fig. 1 and Supplementary Table 3). Only people with ASD showed a higher positive deviation burden than controls ($P < 0.001$; Extended Data Fig. 1 and Supplementary Table 3). Scan quality (for a definition, see Methods) was not correlated with extreme deviation burden (HC_{test}: $\rho = 0.04$, $P = 0.51$; cases: $\rho = 0.03$, $P = 0.34$). In the following, we first focus on characterizing negative GMV deviations (that is, GMV values lower than normative expectations) before considering positive GMV deviations.

Heterogeneity at the level of brain regions

We quantified regional heterogeneity in GMV deviations (Fig. 1, left) as the proportion of individuals showing an extreme deviation in each parcellated brain area, estimated separately for each diagnostic group and the HC_{test} cohort (Fig. 2f). Despite most people showing at least one deviation in each group, the maximum percentage overlap within any of the 1,032 brain regions never exceeded 7% (ADHD: 3.27%; ASD: 4.95%; BP: 5.26%; MDD: 6.21%; OCD: 4.19%; SCZ: 4.96%; HC: 2.97%). Hence, individual extreme deviations were common, but they were rarely found in consistent locations in individuals with the same diagnosis (Extended Data Fig. 2).

We next compared the spatial overlap of each clinical group with controls by subtracting their respective percentage overlap values in each region (Fig. 2g). The statistical significance of the observed Δ overlap values in each region was assessed with respect to an empirical null distribution generated by shuffling group labels (Methods and Extended Data Fig. 3a–d). While each disorder showed isolated areas of greater overlap at uncorrected levels, only ASD (32 regions) and BP (45 regions) showed differences that survived false discovery rate (FDR) correction ($P_{\text{FDR}} < 0.05$, two tailed; Fig. 2h). These differences were scattered throughout the cortex and rarely aggregated into spatially structured clusters. Few regions showed significantly greater overlap in controls compared with cases (Supplementary Fig. 4). Repeating the analyses using a method that avoids reliance on a single threshold for defining extreme deviations³⁰ yielded similar findings (Methods and Supplementary Figs. 5–7). Collectively, these results extend past reports^{10,13,14} to indicate that minimal spatial overlap in the location of person-specific extreme GMV deviations is a general characteristic of psychiatric illness.

Heterogeneity at the level of functional circuits

We next asked whether the regionally heterogeneous extreme deviations identified in each clinical group show significant functional coupling (FC) with common, remote areas, thus yielding greater interindividual consistency at the level of distinct functional circuits (Fig. 1, middle). To this end, we took each region showing an extreme deviation in each participant (Fig. 3a) and mapped its pattern of whole-brain FC in an independent sample of 150 unrelated HCs (HCP₁₅₀) to establish the normative pattern of expected FC for the deviant region (Fig. 3b). We thresholded ($P_{\text{FWE}} < 0.025$) and binarized each deviant-related FC map (Fig. 3c), and took the union of these thresholded maps across all extreme deviations for a given person (Fig. 3d; for a discussion of thresholding issues, see Discussion), yielding a map that represents all areas showing significant FC with at least one extreme deviation in that individual. Next, we estimated, for each region, the proportion of individuals within each group for whom that region showed significant FC with an extreme deviation (Fig. 3e). This analysis revealed that the overlap observed in brain regions functionally coupled to deviant loci was much higher, in absolute terms, than the overlap observed in the locations of the extreme deviations

themselves (Extended Data Fig. 4). For instance, the maximum circuit-level overlap observed across regions was 33% in HC_{test} and ranged between 39% (ADHD) and 53% (SCZ) in the clinical groups.

To a certain extent, one should expect overlap to be higher at the level of FC union maps (Fig. 3d,e) compared with regional deviation maps (Fig. 2e,f) since any single deviant area can show FC with multiple other regions, increasing the likelihood that common areas will be implicated across individuals. For this reason, regional circuit-level overlap maps in each clinical group must be contrasted with the HC_{test} overlap map, which provides a normative benchmark for the expected level of overlap in FC union values (Fig. 3f). A critical consideration in these contrasts concerns the effect of total deviation burden, which differs between groups (for example, Extended Data Fig. 1). For instance, the total number of extreme deviations identified in individuals diagnosed with SCZ was 4,951 compared with only 1,410 in HC_{test}.

This discrepancy means that more deviant-related FC maps will be used in the FC union maps of SCZ individuals, thereby increasing the chance of observing higher overlap. On the one hand, this higher circuit-level overlap will have real phenotypic consequences, since a higher deviation burden is an intrinsic and expected characteristic of psychiatric disorders and these deviations are likely to impact circuit-level function and behavior. On the other hand, it is informative to determine whether the overlap is driven simply by group differences in deviation burden or reflects a preferential targeting of the circuit by the disorder in question. We therefore evaluated the statistical significance of regional group differences in circuit-level overlap using two different permutation tests designed to disentangle these effects (for details, see Methods and Extended Data Fig. 3). The first, group-based permutation test, relied on shuffling the group labels of the individual-specific FC union maps (Extended Data Fig. 3a–d) to assess differences in overall circuit-level overlap regardless of group differences in total deviation burden, thereby characterizing group differences in their ‘natural’ state. The second, spatial permutation test, evaluated group differences with respect to a null distribution that preserves the number of deviation-related FC maps contributed by each individual, thus matching differences in the total deviation burden of each group (Methods and Extended Data Fig. 3e–k). This analysis tests whether observed group differences in circuit-level overlap are greater than expected for FC maps generated from the same number of randomly chosen seeds, with the random seeds being selected from a deviation map with the same underlying spatial autocorrelation as the empirical deviation maps (for further details and interpretation, see Methods). In simple terms, the group-based permutation test identifies naturally occurring group differences in overlap regardless of deviation burden, whereas the spatial permutation test identifies differences in overlap beyond variations in overall deviation burden, implying a preferential targeting of specific functional circuits.

Using the group-based permutation tests, we observed significantly greater overlap across wide swathes of cortex for people with SCZ and MDD ($P_{\text{FDR}} < 0.05$, two tailed) compared with controls (Fig. 3g). Regions with significantly greater overlap in SCZ were distributed diffusely and included ~75% of cortical areas. In MDD, ~31% of areas were implicated, predominately localized to regions of visual, parietal, somatomotor, frontal and insula cortices. A similar spatial pattern was observed in ASD, OCD and BP when considering uncorrected results, although fewer regions survived FDR correction. Very few areas showed greater overlap of deviant-related FC in controls compared with cases (Supplementary Fig. 8a). An alternative method for mapping FC results to parcellated regions yielded comparable findings (Methods and Supplementary Figs. 9, 10a and 11a).

Extended Data Fig. 5 confirms that, for nearly all areas where we identified greater circuit-level overlap in cases compared with controls (Fig. 3g), the magnitude of the difference exceeded the overlap differences observed at the regional level. This result aligns with our main findings to indicate that the level of overlap across participants

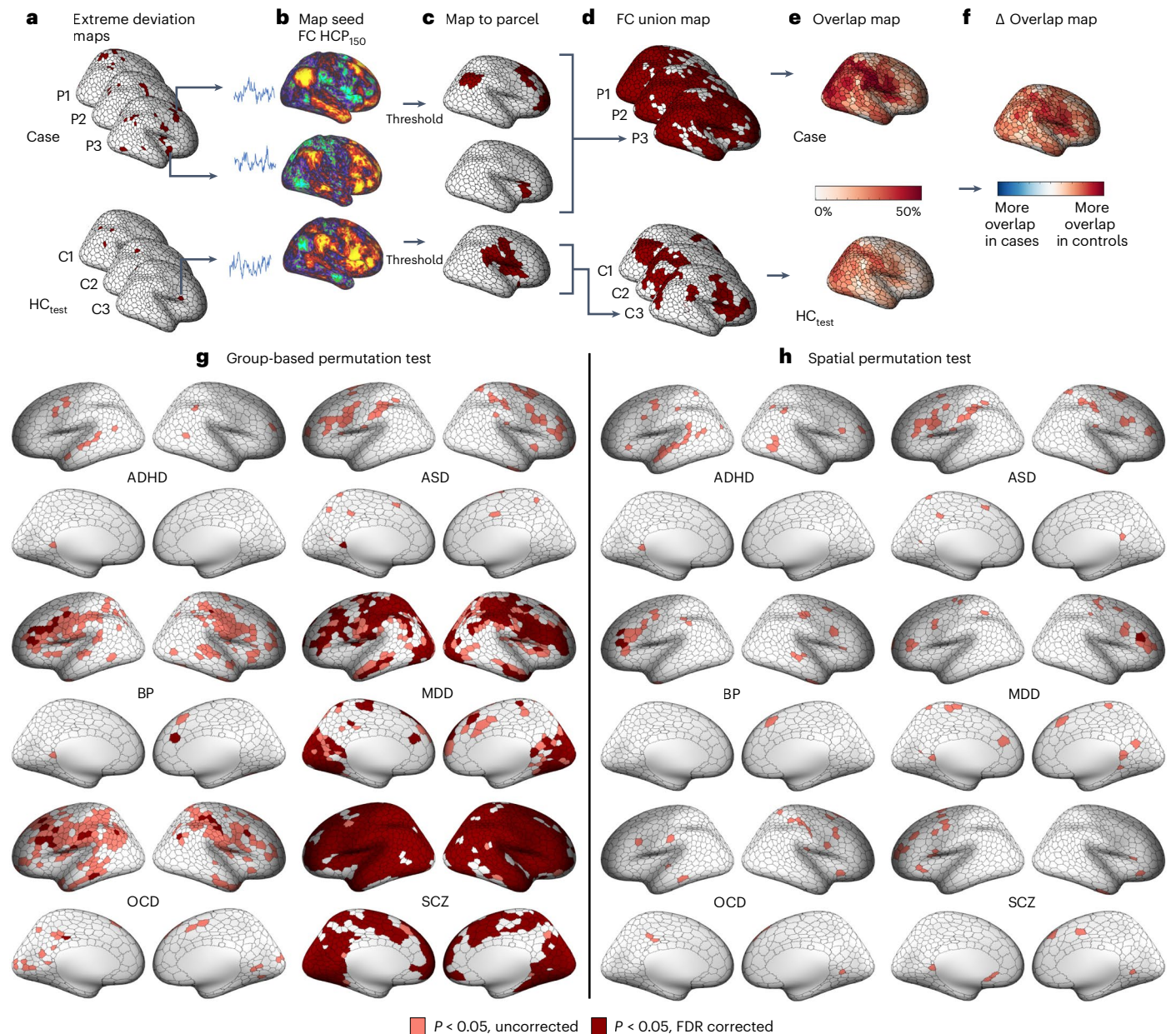


Fig. 3 | Functional circuit heterogeneity of extreme negative GMV deviations in each disorder. a–f, Workflow for characterizing circuit-level heterogeneity. For each participant in the HC_{test} and each clinical group, we took each brain region showing an extreme deviation (a). For each individual in an independent sample of controls (HC_{150}), we extracted a representative time course from each deviant region and mapped the areas to which it is functionally coupled to using a seed-related FC analysis. Shown here are three participants in a clinical group (Case: P1, P2, P3) and three participants in the control group (HC_{test} : C1, C2, C3). Two FC maps for two different deviant loci identified in P3, and one FC map for one deviant loci identified in C3 are depicted (b). We thresholded and binarized each FC map associated with a given extreme deviation (c). Note, that no subcortical regions survived this thresholding procedure. We took the union of the thresholded maps across all deviant FC maps to obtain a single map of all

areas showing direct FC with one or more deviant areas for a given individual (d). For the HC_{test} and each clinical group, we quantified the proportion of individuals showing significant FC in a given region, yielding an extreme deviation FC overlap map (e). We subtracted the HC_{test} FC overlap map from each clinical group's FC overlap map to obtain an FC Δ overlap map for each clinical group. Group differences in circuit-level overlap were evaluated with respect to two empirical null models (for details, see Extended Data Fig. 3) (f). **g, h**, Cortical surface renderings of regions with significantly greater overlap in cases compared with controls in areas functionally coupled to extreme deviations identified using group-based (g) or spatial permutation tests (h) (pink corresponds to $P_{uncorrected} < 0.05$, red corresponds to $P_{FDR} < 0.05$, two tailed, cases > controls). Data used to generate this figure can be found in Supplementary Data 1 (Circuit_neg_parc50).

was greater at the circuit compared with regional level, supporting the conclusion that GMV deviations in all psychiatric disorders except ADHD are part of common functional circuits, despite being located in anatomically heterogeneous areas.

We next used the spatial permutation test to determine the degree to which the above differences in circuit-level overlap were attributable to deviation burden. The results indicate that all disorders show some

evidence of greater overlap in regions of left inferior and middle frontal gyri at uncorrected thresholds (Fig. 3h). However, only three areas in BP and one in MDD survived FDR correction. Few results survived FDR correction when considering regions showing greater overlap in controls (Supplementary Fig. 8b) and an alternative method for mapping FC results to our regional parcellation yielded comparable findings (Methods and Supplementary Figs. 10b and 11b). Thus, while these

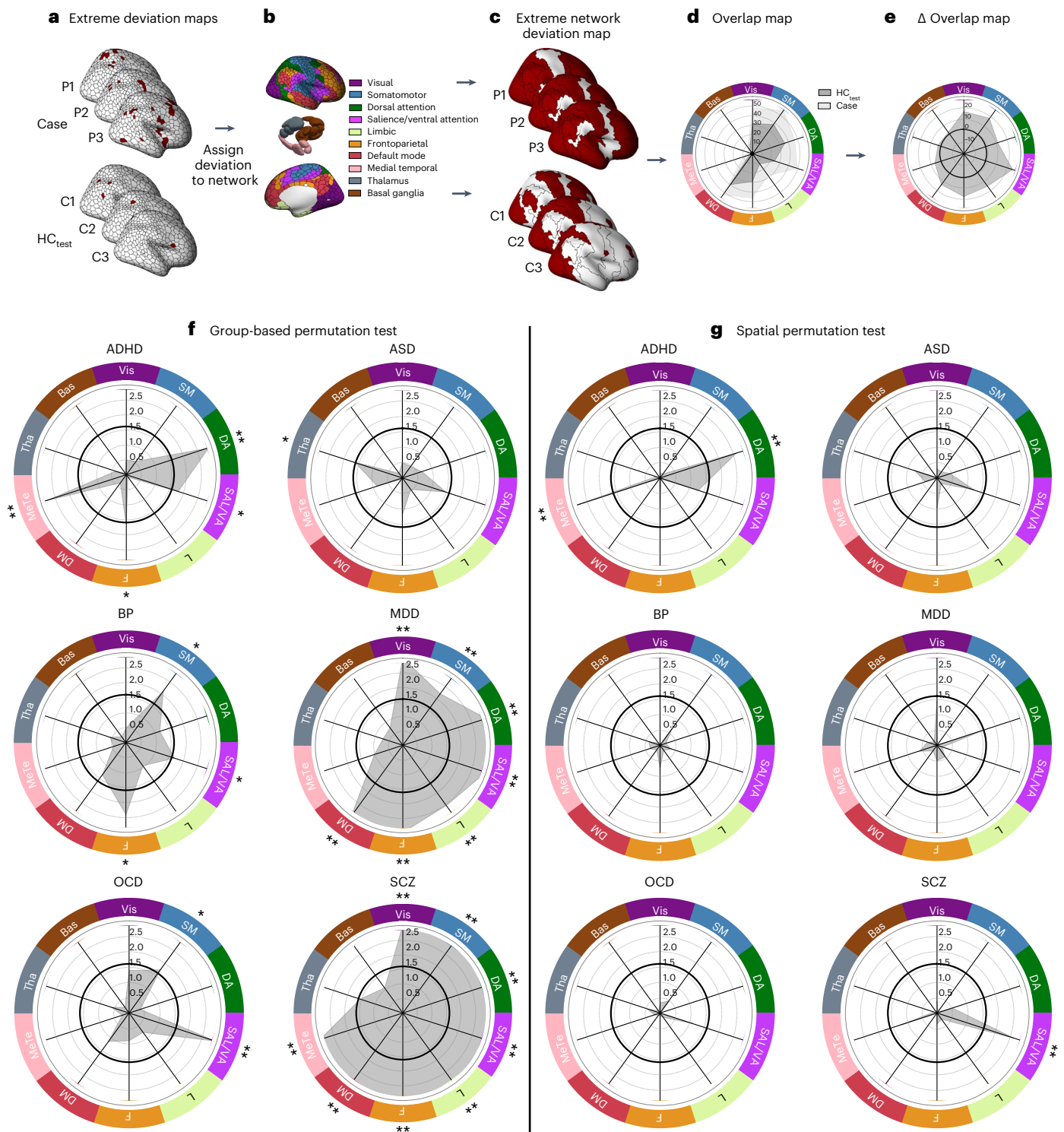


Fig. 4 | Functional network heterogeneity of extreme negative GMV deviations in each disorder. **a–d**, Workflow for characterizing network-level GMV heterogeneity. For each individual in the HC_{test} and each clinical group (**a**), we assigned each brain region showing an extreme deviation to one of seven canonical cortical functional networks or three subcortical nuclei (**b**), such that the entire network was considered deviant if it contained at least one region with an extreme deviation. The cortical surface renderings show the resulting network-level extreme deviation maps (**c**). We quantified the proportion of individuals in each group showing a deviation within each network and compared these proportions to the network overlap in HC_{test} (**d**). **e**, Group differences in network-level overlap were evaluated with respect to two empirical

null models (for details, see Extended Data Fig. 3). **f, g**, The network-level $-\log_{10} P$ values associated with the difference in percent overlap for extreme negative GMV deviations between each clinical group and the HC_{test} cohort (gray) under group-based (**f**) or spatial permutation (**g**) testing, respectively. ** corresponds to $P_{FDR} < 0.05$, two-tailed, cases > controls, * corresponds to $P_{uncorrected} < 0.05$, two-tailed, cases > controls. The solid black line indicates $-\log_{10} P = 1.6$ ($P = 0.05$, two-tailed, uncorrected). VIS, visual; SM, somatomotor; DA, dorsal attention; SAL/VA, salience/ventral attention; L, limbic; F, frontoparietal; DM, default mode; MeTe, medial temporal; Tha, thalamus; Bas, basal ganglia. Data used to generate this figure can be found in Supplementary Data 1 (Network_neg_10network).

findings offer preliminary evidence for the preferential involvement of neural circuitry involving the lateral prefrontal cortex (PFC) in each disorder, the dominant factor explaining greater circuit-level overlap in cases is total deviation burden. In other words, cases show greater overlap at the circuit level largely because they are more likely to express extreme GMV deviations, which in turn increases the probability with which sites functionally coupled to deviant areas will be implicated.

Heterogeneity at the level of extended functional networks

Our analysis thus far indicates that the locations of regional GMV deviations show marked individual heterogeneity, that cases show substantially greater overlap when considering the functional circuitry of these deviant loci and that this overlap is largely driven by total deviation burden. However, our circuit-level analysis only focused on areas showing significant FC with deviant regions, and these circuits often form part of larger, extended functional networks that may not be fully mapped by a circuit-level characterization (Fig. 1, right). We therefore examined the location of person-specific GMV deviations in relation to canonical functional networks, defined according to a widely used and validated classification of brain areas into one of seven cortical networks^{27,31} or three subcortical regions²⁸, to derive a comprehensive, multiscale characterization of GMV heterogeneity (Fig. 4a–c). In this way, if an individual showed an extreme deviation in at least one region affiliated to a given network, the entire network was considered deviant (Fig. 4c). Once again, we quantified the proportion of individuals within each group showing a deviation in each network (Fig. 4d) and compared these proportions between the HC_{test} and each clinical group (Fig. 4e). We then evaluated group differences in network-level overlap using both group-based and spatial permutation tests, as was done in the circuit-level analysis (Extended Data Fig. 3).

The results of these analyses are shown in Fig. 4d–f (for a summary of the degree of spatial overlap (%) in each network for each group, see Supplementary Table 4). Using group-based permutation testing, SCZ and MDD cases showed significantly greater overlap in multiple networks compared with controls ($P_{\text{FDR}} < 0.05$, two tailed; Fig. 4f). For SCZ individuals, the differences included all networks except for the thalamus and basal ganglia. In MDD, all cortical networks were implicated. The dorsal attention network and medial temporal lobe were implicated in ADHD and the ventral attention network in OCD. No networks survived multiple comparison correction for ASD and BP. At uncorrected levels, the salience/ventral attention network was implicated across all disorders except for ASD. There were no networks in which controls showed greater overlap (Supplementary Fig. 12a).

Using spatial permutation tests with FDR correction, only the salience/ventral attention network showed greater overlap in SCZ, and the medial temporal lobe and dorsal attention networks showed greater overlap in ADHD (Fig. 4g). No deviant-related networks showed greater overlap in controls (Supplementary Fig. 12b).

We repeated the same analyses using a 20-network parcellation (17 cortical networks and 3 subcortical nuclei). The results of these analyses are shown in Supplementary Fig. 13 (also see Supplementary Table 5). Broadly, the findings align with those obtained using the ten-network parcellation to indicate that SCZ and MDD are associated with greater network overlap in negative GMV deviations across most networks, with the effects being more circumscribed in the other disorders. Using a 20-network parcellation, we found robust evidence for preferential involvement for temporoparietal and control networks in ADHD and MDD, respectively.

Together, these results align with the circuit-level analysis to indicate that SCZ and MDD are associated with greater network-level overlap of negative GMV deviations. Group differences in overlap are more circumscribed in ADHD, ASD, BP and OCD, although the salience/ventral attention network is implicated across most disorders. These effects are largely driven by total deviation burden, with only SCZ and ADHD showing robust evidence supporting preferential involvement

of the salience/ventral attention and dorsal attention/medial temporal systems, respectively.

Analysis of positive GMV deviations

For completeness, we repeated the same analyses as above for extreme positive GMV deviations, representing areas where volume was higher than normative expectations. At the regional level, extreme deviation overlap never exceeded 6% (ADHD: 5.23%; ASD: 3.96%; BP: 4.82%; MDD: 5.59%; OCD: 4.19%; SCZ: 5.22%; HC_{test}: 2.60%) and very few regions showed significant case–control differences in overlap (Extended Data Figs. 6 and 7 and Supplementary Fig. 14).

Circuit-level overlap was higher, reaching a maximum of 40% across all regions and disorders (Extended Data Fig. 8). Group-based permutation testing revealed significantly greater circuit-level overlap in individuals diagnosed with ASD compared with controls in ~20% of regions, predominately in visual, parietal and frontal cortices. No other differences survived FDR correction (Extended Data Fig. 9a and Supplementary Fig. 15a). Similarly, spatial permutation testing only identified isolated areas in pregenual cingulate and right lateral PFC as showing significantly greater overlap in MDD (Extended Data Fig. 9b and Supplementary Fig. 15b).

At the network level, group overlaps were as high as 48% (Supplementary Table 5), with group-based permutation tests identifying significantly greater overlap in the basal ganglia in SCZ and all cortical networks except the default mode network in ASD ($P < 0.05$, two tailed), compared with controls (Extended Data Fig. 10a). Only the former difference was also observed with spatial permutation tests, and was also accompanied by greater overlap in the salience/ventral attention network (Extended Data Fig. 10b). The medial temporal lobe showed significantly greater overlap in controls compared with SCZ (Supplementary Fig. 16).

In summary, elevated circuit- and network-level overlap was particularly prominent in ASD under group-based permutation testing, and implicated areas of medial and lateral parietal, temporal and PFC at the circuit level, and all cortical systems except the default mode network at the network level. These differences were not apparent with spatial permutation testing, indicating that they were largely driven by the elevated positive GMV burden of individuals diagnosed with ASD. Differences in circuit- and network-level overlap for positive GMV deviations in other disorders were less pronounced.

Discussion

We adapted lesion network mapping^{22,23} for use with normative modeling^{11,12} to characterize individual heterogeneity of brain deviations at the regional, circuit and network level in different psychiatric disorders. We showed that heterogeneity between individuals in regional GMV deviations is a general feature of psychiatric illness, but that these regionally heterogeneous loci are often embedded within common functional circuits and networks. Regional heterogeneity thus offers a plausible explanation for the well-described clinical heterogeneity observed in psychiatric disorders^{8,9}, while circuit- and network-level aggregation of deviations is a putative neural substrate for phenotypic similarities between individuals assigned the same diagnosis. Using different null models, we showed that much of the elevated overlap observed at the circuit and network levels is attributable to total deviation burden. This result challenges prominent models in which specific disorders are assumed to arise from dysfunction of characteristic neural systems selectively targeted by the disease process³².

Patient-specific deviations are regionally heterogeneous

Few areas showed significantly greater regional overlap in cases compared with controls, and no single area in any disorder showed an extreme deviation in more than 7% of cases across the 1,032 regions investigated. This substantial regional heterogeneity aligns with

past normative modeling studies of GMV in SCZ, ADHD and BP^{10,13,14}, confirming that this heterogeneity is a general feature of psychiatric illness. The variable involvement of different regions across individuals may yield distinct clinical profiles and drive phenotypic heterogeneity in people with the same diagnosis^{8,9}. An important avenue of future work will involve precisely characterizing the relationship between GMV deviations and interindividual differences in symptom expression.

More generally, the high degree of regional heterogeneity observed indicates that group mean comparisons are not representative of the specific profile of GMV deviations apparent in any individual case. Mean comparisons may thus offer an incomplete account of causal pathophysiological mechanisms, unless the broader network context of any identified group differences is considered. Notably, most participants showed a low deviation burden (Supplementary Table 3). This result is expected, since the failure so far of psychiatric neuroimaging to identify robust diagnostic biomarkers of illness indicates that any disease-related brain changes are likely to be subtle and complex. We should, therefore, be circumspect about our ability to identify strong neurobiological signatures of psychiatric disorders.

Deviations aggregate in common circuits and networks

Despite considerable heterogeneity at the regional level, deviations were often coupled to common functional circuits and networks. In some cases, >50% of people with the same diagnosis showed a deviation implicating at least one of these systems, with group-based permutation testing indicating that all disorders showed some evidence of greater circuit-level overlap than controls. This higher overlap at circuit and network levels parallels lesion network mapping studies of neurological syndromes, which suggest that many clinical phenotypes are not caused by dysfunction of the lesioned area itself but by its impact on remote, functionally coupled regions^{22–24}. Our findings indicate that a similar process may occur in psychiatric disorders, with anatomically distributed GMV deviations often being coupled within similar circuits and networks. The consequent dysfunction of these common circuits and networks may drive clinical similarities between people with the same diagnostic label, despite extreme heterogeneity in the locations of the deviations themselves. Multiple mechanisms may explain these circuit- and network-level effects, ranging from transient diaschisis-like effects on distributed circuit/network function through to more prolonged transneuronal dysfunction caused by aberrant interregional signaling or disrupted axonal transport of trophic factors²⁰.

Areas of frontal, parietal, insula and temporal cortex showed greater circuit-level overlap across most conditions, and cross-disorder differences were more of degree rather than kind. For instance, differences in circuit-level overlap were spatially circumscribed in ADHD but encompassed nearly the entire brain in SCZ (Fig. 3g). These findings challenge classical views that distinct psychiatric diagnoses are associated with dysfunction in specific circuits (for example, ref. 32,) and suggest that each disorder is associated with complex changes that affect diverse neural systems^{33,34}, often transdiagnostically^{35–37}. Accordingly, our network-level analysis revealed greater overlap in the salience/ventral attention network in five of the six disorders that we considered (Fig. 4f). The salience/ventral attention network plays a central role in cognitive control³⁸, interoceptive awareness³⁹, and switching between internally and externally-focused attention⁴⁰. Its dysfunction has been implicated in a diverse range of psychiatric disorders^{41–44}, is associated with increased levels of general psychopathology in youth⁴⁵ and shows cross-disorder abnormalities in meta-analyses of classical case-control voxel-based morphometry (VBM) studies and functional neuroimaging studies^{35–37}. Our findings support this past work to suggest that salience/ventral attention network dysfunction may play a critical role in the expression of general psychopathological processes common to diverse diagnoses⁴⁶.

Generic and preferential targeting of neural systems

Our spatial permutation test allowed us to evaluate the degree to which circuit- and network-level overlap was explained by group differences in total deviation burden. At the circuit level, prefrontal regions were identified as showing greater overlap compared with controls in all clinical groups at uncorrected levels, consistent with extensive evidence for prefrontal dysfunction in each condition^{47–49}. However, only regions of right and left lateral PFC in MDD and BP, respectively, survived FDR correction. The right PFC showed 21% overlap in MDD cases, whereas the left PFC showed 14–15% overlap in BP. Extensive literature implicates lateral PFC dysfunction in affective disorders^{48,50,51}, and the dorsolateral PFC is a popular brain stimulation target in both MDD⁵² and BP⁵³. The region has also been implicated by lesion network mapping as a core site explaining the emergence of depression following stroke⁵⁴. Given that our spatial permutation test did not find any evidence of greater network-level overlap in MDD and BP, our findings point to a strong specificity for functional circuits coupled to these particular prefrontal areas. Spatial permutation testing also revealed greater network-level overlap for SCZ in the salience/ventral attention network and for ADHD in the dorsal attention network and medial temporal regions, suggesting that GMV deviations preferentially target these systems in a disorder-specific way.

Lesion network mapping of neurological cases indicates that symptom expression is often driven by dysfunction of remote sites connected to a lesion, implying that these remote sites represent viable treatment targets^{22,23,55,56}. Following this logic, the findings from our spatial permutation analysis imply that viable treatment targets may be located in the right PFC for MDD, in the left PFC for BP, in the salience/ventral attention network for SCZ, and in the dorsal attention and medial temporal networks for ADHD. However, these targets may only be relevant for a subset of patients, with the observed overlaps ranging between 10% and 50% across disorders. It thus follows that the most current approaches, which attempt to identify a single common therapeutic target for each diagnosis, will only have limited success. A more comprehensive understanding of patient-specific brain changes, and their network context, will be necessary to develop more effective, and personally tailored, interventions.

In most other cases, we failed to reject the spatial null hypothesis. This result indicates that many case-control differences in circuit/network overlap observed under group-based permutation cannot be attributed to the preferential accumulation of deviations within a particular circuit/network, since the differences are consistent with comparisons of the same number of randomly selected seeds in each group. When taken with our group-based permutation analysis, our findings indicate that accumulating deviations are more likely to be coupled to areas in prefrontal, temporal, parietal and insula cortices, simply because these areas are known connectivity hubs of the brain and have a higher probability of being functionally coupled to other regions^{57–59}. While this circuit- and network-level accumulation will still have real phenotypic consequences, the spatial permutation test allowed us to identify an underlying random process as a candidate generative mechanism. This finding challenges the implicit assumption in many studies that any brain changes observed in a disorder result from a targeted pathophysiological process. Further work investigating the genetic and environmental contributions to individual-specific deviations should elucidate the mechanisms driving their anatomical distribution.

Heterogeneity of positive deviations

The locations of positive GMV deviations were more heterogeneous than negative deviations, and showed less overlap across people. The exceptions were ASD and BP at the regional level, which showed significantly greater overlap in predominately frontal, parietal areas. At the circuit level, ASD showed significantly greater overlap in visual, parietal and frontal cortices, and at the network level, all cortical networks

except the default mode network were implicated in ASD. ASD has been associated with dysregulated and accelerated brain growth, particularly in the temporal, parietal and frontal association cortices during early childhood⁶⁰. Whether these increases persist into adulthood and can explain the present findings remains unclear.

Differences in circuit-level and network-level overlap for positive GMV deviations in other disorders were less consistent. In MDD, greater circuit-level overlap in left pregenual cingulate and anterior right lateral PFC is consistent with the known roles of these areas in regulating emotion⁶¹ and cognitive control⁶², respectively. The greater overlap observed in the basal ganglia of people with SCZ may be attributable to the effects of antipsychotics, which can cause volumetric expansion in this region⁶³.

Limitations

While deviation burden was higher in cases than controls, and most people showed at least one deviation, a substantial fraction of clinical participants (~50%) showed a relatively low deviation burden (<3 deviations). This result implies that there is considerable overlap between cases and controls, as noted for many other phenotypes⁶⁴, and that brain changes associated with psychiatric illness may be subtle and complex. This subtlety, combined with considerable interindividual heterogeneity, may partly explain the failure of the field to identify pathognomonic biomarkers for psychiatry, despite decades of research. The result may also reflect the limited sensitivity of MRI-derived GMV estimates to map pathophysiologically relevant brain changes. Normative modeling of other phenotypes, such as those obtained with functional or molecular imaging, may reveal stronger separation between groups.

Sample sizes in single-site psychiatric neuroimaging studies are often small, so we pooled data from multiple sites to generate a sufficiently large cross-disorder dataset. As a result, the data were collected with different acquisition, recruitment and clinical assessment protocols. To avoid introducing a dependence between scan site and diagnostic group, which can confound case-control comparisons, we focused primarily on control data obtained on the same scanners as the clinical data. While our model diagnostics indicated adequate fit to the data, future work could extend the techniques used here so that they can be used with normative models trained on larger samples (for example, refs. 65,66) to obtain more reliable deviation estimates.

Although we used stringent quality control, and our hierarchical Bayesian model²⁹ appropriately parsed site-related variance (Supplementary Table 2), investigating correlations with symptom profiles or other clinically relevant variables such as age of onset, illness duration, medication exposure or disease severity across disorders was beyond our scope. Indeed, while our methods are useful for characterizing circuit- and network-level overlap across individuals, they are not well suited for analyses of individual differences in behavior. A valuable extension of our work would involve directly testing the hypothesis that circuit- and network-level overlap relate to clinical similarities between patients. Addressing this goal would require appropriate methods for aggregating deviation-related circuit- and network-level data within and between individuals, and extensive measurement of diverse symptom dimensions to allow adequate assessment of clinical phenotypes⁶⁷. Extensive symptom sampling that allows for personalized models would be particularly useful in this context⁶⁸. Such data are rarely available in single-site clinical neuroimaging studies. Future studies may work toward developing harmonized, transdiagnostic and multisite clinical protocols that can be used across different disorders, such as those informed by the Hierarchical Taxonomy of Psychopathology⁶⁹. The specific age at which the measures are acquired should also be considered, given that the nature and magnitude of case-control differences may vary across the lifespan⁴. Such an approach may facilitate data-driven strategies to identify biological subtypes based on patient-specific deviation maps that cut across traditional diagnostic boundaries.

Our cases were diagnosed according to Diagnostic Statistical Manual of Mental Disorders (DSM) or the International Classification of Diseases (ICD) criteria^{70,71}. Given the widespread application of DSM and ICD in clinical and research contexts, we deemed it important to understand heterogeneity with respect to these constructs and to gain insight into the neural correlates of phenotypic similarities and differences between cases assigned specific diagnoses. Nonetheless, focusing on specific syndromes rather than traditional diagnoses, as in the lesion network mapping literature^{23,24}, may yield a more precise mapping between regional deviations, their network context and behavior. Establishing sufficiently large databases for conducting such syndrome-focused analyses will be a key challenge for future work.

Conclusions

Our multiscale analysis of neural heterogeneity across six psychiatric disorders confirms that extreme regional heterogeneity of GMV deviations is a general characteristic of mental illness. Further, we showed that these deviations are often coupled to common functional circuits and networks, offering a putative neural substrate for phenotypic similarities among individuals assigned the same diagnosis. The common involvement of prefrontal and parietal circuits, and the salience/ventral attention network, across disorders may be a marker of transdiagnostic psychological distress, with the variable involvement of other systems explaining phenotypic differences among disorders. More broadly, our findings underscore the need to consider the network context of disorder-related markers of pathophysiology²⁰, and indicate that the clinical expression of disease is not driven solely by sites of primary pathology, but also by the effect of this pathology on remote, connected systems.

Mapping neural heterogeneity at the regional level

Our analysis aimed to characterize neural heterogeneity at the level of individual brain regions, neural circuits and extended brain networks (Fig. 1). To map heterogeneity at the regional level, we obtained person-specific deviation maps, representing the extent to which the GMV of a given person within each brain region deviates from normative predictions. This analysis involved quantifying regional GMV for each individual and evaluating the observed measures with respect to an underlying normative model, as outlined in the following.

Anatomical data

We estimated regional GMV using VBM of T1-weighted anatomical MRI scans^{25,26}. We focused on GMV because it is one of the most frequently studied neural phenotypes in psychiatry. VBM is also arguably the most widely used tool for measuring GMV in this context. Image acquisition parameters for each dataset are provided in Supplementary Table 1. The following outlines our quality assurance procedures and data processing pipeline, which was applied to all raw T1-weighted images obtained for each dataset.

Quality control

All T1-weighted images were visually inspected and evaluated for the presence of artifacts^{72,73}, resulting in the exclusion of 53 images with gross artifacts or abnormalities. Next, we used the Computational Anatomy Toolbox²⁵ (CAT12 r1113, ref. 74) to generate a weighted overall image quality rating (IQR) for every scan. This metric combines ratings of basic image properties, including the level of noise and geometric distortions, into a single score that quantifies the overall image quality of a participant's T1-weighted scan (for more information, see ref. 74). On this metric, lower scores denote higher image quality. As per previous work¹⁰, we excluded 153 images with an IQR >2.8. An additional 63 images were excluded due to a failure of our processing pipeline. Collectively, this quality assurance process resulted in the exclusion of a total of 269 scans

MRI preprocessing

Regional GMV was estimated using the CAT12 VBM pipeline^{25,26}, which is included as an extension of Statistical Parametric Mapping software v7771²⁵ in MATLAB v9.8 (ref. 76). Briefly, the T1-weighted images were first corrected for intensity nonuniformities, and segmented into gray matter (GM), white matter and cerebrospinal fluid tissue probability maps. Then, using the high-dimensional Diffeomorphic Anatomical Registration Exponentiated Lie Algebra⁷⁷, the segmented scans were normalized into standard Montreal Neurological Institute Ix1555 space. Lastly, the images were bias-field corrected and modulated by the linear and nonlinear components of the Jacobian determinant obtained from the Diffeomorphic Anatomical Registration Exponentiated Lie Algebra deformation fields to obtain voxel-wise estimates of GMV. To constrain our analyses to GM voxels, we generated a mean image from all the normalized GM maps and retained voxels with a tissue probability ≥ 0.2 .

Brain parcellation

We employed HBR to estimate normative models of regional GMV (see the Normative Model section, below). To limit computational burden, we parcellated the brain into 1,032 cortical and subcortical regions by combining well-validated parcellations of the cortex²⁷ comprising 1,000 regions, and of the subcortex²⁸ comprising 32 regions. The cortical parcels have been mapped to well-described, canonical functional networks of the brain³¹, facilitating our analysis of network-level overlap in deviations.

Recent work indicates that group-based parcellations can be tailored to capture individual variability in regional borders (for example, ref. 78). We used a group-based parcellation to ensure a consistent approach across cortex and subcortex (methods for individual subcortical parcellations have not yet been developed), to allow direct comparison with extant literature (which have overwhelmingly relied on group parcellations), and to ensure that region-of-interest seeds remained approximately the same size among participants, as the effect of varying seed sizes between individuals has not yet been studied extensively. Investigating the effect of individually tailored parcellations on our findings will be an important goal for future work. Voxels that overlapped between atlases were assigned to the corresponding subcortical region. Regional GMV estimates were obtained using freely available code⁷⁹.

The choice of parcellation will necessarily affect the degree of overlap observable in any given brain area, such that coarser parcellations will be associated with higher levels of overlap. Comparison with the control data thus provides a critical normative benchmark against which to evaluate the levels of overlap observed in each clinical group. We chose our 1,032-region parcellation to offer high spatial resolution while ensuring computational feasibility. The same parcellation was used in regional-, circuit- and network-level analyses, facilitating direct comparison across scales.

Normative modeling

We used normative modeling to obtain person-specific GMV deviation maps in relation to an underlying model of normative expectations for regional GMV variations (see refs. 11,12,29, PCNtoolkit v0.16 (ref. 80)). Normative models estimate the mean and variance (referred to as a normative range) of a response variable (for example, GMV) from a set of clinically relevant covariates (for example, age and sex) across a large healthy sample, referred to as a training set. These estimates are then used to quantify the deviations of samples in the test subset, which typically consists of cases sampled from the normative demographic range of the training set.

When using multisite data, scanner- and site-related variability introduce artefactual variance that confounds the results of any subsequent analyses⁸¹. These confounding effects present as site-correlated biases that cannot be explained by biological heterogeneity between

samples. To account for these site-related effects, we built our normative model using HBR, which successfully accommodates signal and noise variance in multisite data by estimating different but connected mean and variance components through shared prior distributions across sites²⁹. Before modeling, regional GMV estimates were first subjected to a Box-Cox transformation to ensure normality⁸². The optimal lambda parameter for minimizing skewness was estimated for each brain region independently using maximum likelihood.

For the model to accurately parse variance attributable to age, sex and site, sufficient observations are required for any given combination of these variables. We therefore excluded data cells containing less than ten HCs for a particular sex at a given site. If this exclusion procedure resulted in less than ten cases in total from any given site, the entire site was excluded, resulting in the exclusion of 217 scans.

The training data for the normative model (HC_{train}) were created by randomly selecting 90% of HC individuals from each site, provided that the site included data for ≥ 30 HCs. For sites with smaller HC samples, all HC data were included in the training set. The test data (HC_{test}), which was completely independent of the normative model, comprised the 10% of HCs from sites with ≥ 30 HCs and all case scans. The HC_{test} data, comprising 269 individuals (140 male), offered a normative benchmark for assessing case-specific model deviations, as outlined below.

Following stratification of our sample into training and test subsets, for each of the 1,032 parcellated brain regions, we fitted separate HBRs²⁹. The HBRs modelled site and sex effects with random slopes, intercepts and noise, to model GMV as a function of age, sex and site in the training data, yielding estimates of normative regional GMV variance and predictive uncertainty. Importantly, based on a partial-pooling approach, shared prior distributions were imposed over site-specific and sex-specific model parameters. These shared priors assume that, while the model parameters for each site and sex are different, they are drawn from a common distribution. These shared priors thus regularize the model parameters and prevent the model from overfitting small batches for a given sex or site.

To quantify deviations from the normative model predictions for each participant in the test data, we generated deviation z maps for each participant. Specifically, for each participant i , at each brain region j , we combined the predicted GMV \hat{y}_{ij} , true GMV y_{ij} , the predictive uncertainty σ_{ij} and the normative variance σ_{nj} to calculate a z score, z_{ij} , as

$$z_{ij} = \frac{y_{ij} - \hat{y}_{ij}}{\sqrt{\sigma_{ij}^2 + \sigma_{nj}^2}},$$

which quantifies the extent to which an individual's regional GMV estimate deviates from the model prediction, given the uncertainty of the model.

For any given brain region, we are interested in participant values that show large deviations from the normative expectations set by the model, under the assumption that these deviations represent pathophysiologically relevant features of the disease¹¹. We used two complementary approaches to characterize large positive and negative deviations from the normative model at each region and for each individual. First, we thresholded the deviation maps at $z > |2.6|$ (that is, $P < 0.005$) to identify regions showing extreme deviations from model predictions. These deviation maps were derived for each individual in the clinical and the held-out control (HC_{test}) groups and were used for subsequent analyses. We used this approach to be consistent with prior work^{10,14} and because alternative methods, such as the FDR⁸³, rely on adaptive thresholds that can vary between individuals. Applying a fixed z threshold allows for an absolute definition of extreme deviations. Second, to ensure that our findings were not driven by this specific choice of threshold, we adapted a threshold-weighted approach³⁰, as detailed below.

Note that scaling deviation z scores by predictive uncertainty means that we are less likely to observe deviations in data regimes where model uncertainty is high. This may be a consideration for deviation estimates in older participants, given that our training sample was skewed toward younger ages. However, we found no correlation between age and deviation burden ($\rho = -0.004, P = 0.86$).

Evaluation of model performance

To assess model generalizability, we used five-fold cross-validation applied to the HC_{train} cohort ($n = 1,196$ HC individuals, 55% male). Specifically, we partitioned the HC_{train} cohort into five folds. Within each fold, we trained HBR models on 80% of participants, using age, sex and site as covariates, withholding 20% of the participants for estimating generalization performance. This procedure was repeated five times so that regional GMV values for all participants in the HC_{train} group were predicted once. As above, we estimated deviation z maps for every individual and identified extreme deviations. This procedure is standard in machine learning and provides approximately unbiased estimates of the true generalization ability.

For each fold in the cross-validation, we assessed model fit for each brain region by evaluating three performance metrics: (1) explained variance, (2) the mean standardized log-loss and (3) the standardized mean-squared error²⁹, as shown in Supplementary Fig. 3. We also evaluated the model's efficacy in partitioning site-related variance in the data using linear support vector machines (LSVMs). Specifically, we used a series of one-versus-all LSVMs (with default slack parameter of 1) trained separately on the z maps from the HC_{train} and the HC_{test} subsets to classify scan sites. For each site, we ran a two-fold LSVM classifier to obtain the mean balanced accuracy score for the given site. Here, a balanced accuracy at chance level (50%) indicated that the resulting deviations were not contaminated by residual site effects, as confirmed in Supplementary Table 2. To assess whether individual variations in scan quality affected the deviation z maps, we calculated the Pearson's correlation between the total number of extrema and CAT12's IQR rating in each group.

Characterizing regional heterogeneity of extreme deviations

We characterized the heterogeneity of case-specific thresholded deviation maps for each disorder using a nonparametric approach. Specifically, for each clinical group and the HC_{test} cohort, we computed the proportion of individuals in each group showing an extreme deviation within each region, resulting in group-specific overlap maps estimated separately for positive and negative extrema (Fig. 2f). We used the proportion of individuals rather than raw counts to account for sample size differences between groups. Next, we subtracted the HC_{test} overlap map from each disorder's overlap map, resulting in an overlap difference map for each disorder (Fig. 2g). We then permuted group labels (that is, HC and case) and repeated the procedure 10,000 times to derive an empirical distribution of overlap difference maps under the null hypothesis of random group assignment (Extended Data Fig. 3a–c). For each brain region, we obtained P values as the proportion of null values that exceeded the observed difference (Extended Data Fig. 3d). The tails of the null distribution (that is, values associated with $P < 0.10$) were approximated using a generalized Pareto distribution⁸⁴, as implemented in the Permutation Analysis of Linear Models software package⁸⁵ (alpha116), to allow inference at arbitrarily high levels of precision. Statistically significant effects were identified using an FDR-corrected⁸³ threshold of $P_{\text{FDR}} < 0.05$, two tailed.

Threshold-weighted deviation mapping

The procedure to characterize the regional heterogeneity of extreme deviations above used an arbitrary threshold ($z < |2.6|$) on each individual z map. The practical benefit of this approach is that it allows calculation of intuitive metrics, such as the proportion of cases within a group showing a supra-threshold deviation within a given brain region.

We complemented this analysis with an alternative approach, which does not yield similarly intuitive metrics but which integrates results across a range of thresholds³⁰, thus allowing us to determine the degree to which our findings depend on a specific threshold choice. First, for each diagnostic group, we thresholded z maps across the threshold range $1.64 < z < 3.10$, in 100 equal log-space increments, separately for positive and negative extrema. Second, we obtained a region- and threshold-specific percentage overlap map quantifying the proportion of individuals within each group showing an extreme deviation for any given region and threshold. Third, we applied a weighted function to penalize less conservative thresholds. Fourth, we obtained a final threshold-weighted overlap map by taking the area under the curve of the cumulative histogram of the proportion of participants at each threshold for each region across participants. We used the following linear weighted function proposed by Seghier and Price³⁰, which ensures that overlap map values range between 0 and 1, with 1 indicating an effect in a region is present in each participant at each threshold:

$$W_{\text{th}} = 2 \times \frac{\text{th} - z_{\text{min}}}{z_{\text{max}} - z_{\text{min}}},$$

where z_{max} and z_{min} respectively correspond to the maximum and minimum z -value thresholds, and th is the given statistical threshold. We then calculated the difference between the threshold-weighted percentage overlaps between cases and the held-out HC cohort, resulting in an overlap difference map for each disorder. We performed inference on these differences using the same group-based permutation procedure described above with statistical significance defined using a threshold of $P_{\text{FDR}} < 0.05$, two tailed.

Mapping neural heterogeneity at the circuit level

Having investigated neural heterogeneity at the level of individual brain regions, we next evaluated whether regions showing extreme deviations in individual cases are functionally coupled to common areas, which we term circuit-level heterogeneity (Fig. 1, middle). Where prior work has identified spatial correlations between GMV difference maps and canonical functional networks estimated at the level of group means⁸⁶, here we leverage normative modeling to move beyond such means to investigate whether person-specific GMV deviations aggregate within common functional circuits. Evidence for such aggregation offers a plausible explanation for phenotypic similarities between individuals assigned the same diagnosis, despite reports of extreme heterogeneity in deviation location at the regional level^{10,12–18}. We define the functional circuitry of a brain region showing an extreme deviation as the set of brain areas that show significant FC with that region in an independent cohort of healthy individuals. To map this circuitry, we adapted elements of the lesion network mapping methodology used extensively in neurological disorders^{22–24} for use with normative modeling, as outlined in the following.

fMRI data and processing

The functional circuitry of deviant regions was mapped using an independent cohort of 150 resting-state functional MRI (rs-fMRI) scans (71 males, age 21–35 years) from the S900 release of the Human Connectome Project⁸⁷ (HCP). These individuals corresponded to the 150 people with the lowest total head motion, as estimated using frame-wise displacement⁸⁸, from the broader set of 282 unrelated participants with the same fMRI reconstruction in the S900 release. We focused on this subsample to minimize the effects of head motion in our FC maps and to minimize computational burden due to the large number of analyses required. Mapping deviation-related functional circuitry in an independent cohort is essential because we seek to understand the circuit context of deviant regions under normative conditions. Our approach aligns precisely with the logic used in lesion network mapping of neurological patients^{22,23}.

The rs-fMRI data for each participant consists of four runs (rfMRI_REST1_LR, rfMRI_REST1_RL, rfMRI_REST2_LR and rfMRI_REST2_RL) acquired at two different sessions (REST1 and REST2) using two different directions of phase coding (LR: left to right, and RL: right to left). Here, we mapped deviation-related FC using the rfMRI_REST1_LR data. All rs-fMRI data were obtained using a 32-channel Siemens 3T connectome-Skyra scanner. The imaging parameters for rs-fMRI were as follows: repetition time 720 ms, echo time 33.1 ms, flip angle 52°, field of view 208 × 180 mm², matrix 104 × 90, slice number 72, slice thickness 2 mm, voxel size 2 × 2 × 2 mm³, multiband factor 8 and 1,200 volumes. Participants were required to stay awake, relaxed and to keep their eyes open and fixed on a bright cross-hair projected on a dark background. A detailed description of the HCP data is available in previous work⁸⁷.

The rs-fMRI data underwent the HCP's minimal preprocessing pipeline⁸⁹, which includes gradient-nonlinearity-induced distortion, motion correction to the single-band reference image using FMRIB's Linear Image Registration Tool (FLIRT), echo-planar imaging (EPI) image distortion correction using TOPUP⁹⁰; registration into standard space using a customized boundary-based-registration algorithm and single-step spline interpolation using all transforms; intensity normalization and bias field removal to resample the original EPI into Montreal Neurological Institute space. Minimal high-pass filtering was applied with a cutoff of 2,000 ms. Artifacts were then removed using independent component analysis X-noiseifier (ICA-FIX)⁹¹. This involves employing an automatic classifier, specifically trained for HCP data, to identify ICA components due to measurement noise, additional motion or physiological artifacts such as cardiac pulsation and respiration. Next, the volume time series were mapped into the standard CIFTI grayordinate space⁸⁷ and smoothed to 2 mm full width at half maximum (where the smoothing was on the surface for the cortex and in volume space for subcortex). This approach results in a standard set of grayordinates in every participant that encompasses surface vertex data and subcortical volume voxel data. The mean GM signal was then removed from each grayordinate's time series to remove residual widespread signal deflections that were not removed by ICA-FIX^{92–94}. Global signal regression (GSR) has been criticized because the global signal, by construction, contains neural signal and its removal shifts the distribution of correlations so that it is approximately centered on zero^{95,96}. Nonetheless, it is a highly effective denoising procedure, successfully removing contributions to fMRI signals from head motion and respiratory variations^{97,98}. Invasive recordings in animal models indicate that GSR increases the correspondence between hemodynamic signals and neuronal activity⁹⁹, and FC estimates obtained after GSR in humans are more predictive of a wide range of behavioral measures than FC estimates obtained without GSR¹⁰⁰. GSR also dramatically enhances the anatomical specificity of seed-based correlation maps¹⁰¹, which is essential for precise circuit mapping. We therefore used GSR in the present analysis to enable direct comparison with the vast majority of seed-based FC studies in the literature. Alternative strategies for denoising (for example, refs. 92,94) may address the limitations of GSR and may provide more refined network maps.

Mapping the functional circuitry of extreme deviations

We used seed-based FC analysis to map the functional circuitry of each region showing an extreme deviation in any individual within the clinical and HC_{test} groups. Specifically, each deviant region was used as a seed, from which the average time course was extracted for each individual in the HCP₁₅₀ sample (Fig. 4b). These seed time courses were then correlated with all other brain grayordinates and the resulting correlation maps were subjected to Fisher's *r*-to-*z* transformation¹⁰². The transformed correlation maps for each HCP₁₅₀ participant were then aggregated using a one-way *t*-test at each grayordinate, as implemented in Permutation Analysis of Linear Models. Thresholded maps representing the functional circuitry of the seed were obtained using permutation testing (500 permutations) with a generalized Pareto

approximation of the tail of the null distribution⁸⁴ and threshold-free cluster enhancement (TFCE)¹⁰³, run separately for the cortical and subcortical areas, with a threshold of $P_{FWE} < 0.025$, family-wise error corrected (that is, 0.05/2 to account for cortical and subcortical analyses). This procedure was repeated for each region in which at least one extreme deviation across individuals was observed (Fig. 3b).

To quantify circuit-level overlap, we developed an approach that parallels traditional lesion network mapping methodology^{22,23} but which allows statistical inference on observed case–control differences in overlap scores at each region. Specifically, we thresholded each deviant-related FC map by applying TFCE to the grayordinate maps, as described above. We then classified an area in our 1,032-region parcellation as showing significant FC with the seed if more than 50% of its grayordinates survived the TFCE threshold. We repeated our analysis using a mapping threshold of 75% to ensure that our findings were not driven by this specific choice (Supplementary Figs. 9–11). Note, the fsLR32k projection of the Schaefer cortical parcellation is missing two parcels (533 and 903, for more details, see ref. 104) so these were excluded from the analysis.

We mapped FC at the grayordinate level and thresholded the data in this way to leverage the superior statistical sensitivity of TFCE and to map the spatial architecture of the seed-related FC patterns more accurately. Our procedure resulted in a binary map representing the specific brain regions that comprise the functional circuitry of each seed (Fig. 3c). Then, for each individual, we took the union of the binary FC maps across that person's set of extrema, resulting in a single map identifying areas showing significant FC to any extreme deviation expressed by the person (Fig. 3d). Separate union maps were obtained for positive and negative deviations. The overlap of these union maps was then taken for each clinical group and the HC_{test} sample, representing, for each region, the proportion of individuals in that group for whom significant deviant-related FC was identified. We refer to this image as a circuit-level overlap map (Fig. 3e).

Here we mapped the network context of GMV deviations using FC estimates of normative brain architecture mapped in an independent sample, following the approach typically used in lesion network mapping^{22,23}. It is possible that patterns of FC are altered in diagnosed individuals, and that the underlying network architecture differs from normative expectations. The net effect of such alterations would be to affect coupling between deviant and other areas, and thus our normative benchmark still offers an important reference point for understanding the distributed effects of regional deviations. Further investigation of how altered patterns of FC may moderate these relationships remains an important topic for future investigation. Alternative approaches could involve mapping the network context of GMV deviations with estimates of structural connectivity obtained with diffusion MRI^{105,106} or estimates of interregional structural covariance¹⁰⁷.

Characterizing circuit-level heterogeneity

We evaluated group differences in the circuit-level overlap maps of each clinical group relative to the HC_{test} group in two ways (for a schematic overview, see Extended Data Fig. 3). First, following the analysis of regional overlaps, we computed the difference in the case and control overlap maps (Δ overlap map, Fig. 3e) and evaluated the statistical significance of these differences with respect to an empirical null distribution generated by permuting the group labels of the individual-specific FC union maps 10,000 times, with a generalized Pareto tail approximation⁸⁴. Statistically significant differences in deviation FC overlap were identified using a threshold of $P_{FDR} < 0.05$, two tailed (Extended Data Fig. 3a–d).

This group-based permutation test identifies regions where the overlap in deviant-related functional circuitry differs between cases and controls, but it does not preserve group differences in the total number of deviations. It therefore cannot distinguish whether any observed differences arise from a preferential aggregation of

deviations within specific circuits in one group relative to the other, or whether the group differences are driven by variations in total deviation burden. Note that differences in deviation burden should not be thought of as a confound in this context, since they are an intrinsic feature of the disease that will have real phenotypic consequences. Nonetheless, it is important to determine which findings might be driven by differences in deviation burden compared with a preferential accumulation of pathology within a specific circuit for a given disorder.

To disentangle these possibilities, we considered a second, spatial null model that preserves group differences in deviation burden. Specifically, we generated an ensemble of null cortical deviation maps for each individual in the test data by spatially rotating their empirical, unthresholded deviation maps (Extended Data Fig. 3e) using Hungarian spherical spin tests^{108,109} (Extended Data Fig. 3f) and then thresholding these rotated maps in the same way as the observed data (that is, $z > |2.6|$; Extended Data Fig. 3g). The Hungarian method was used as it preserves the original values and spatial autocorrelation of the original map, which ensures that the surrogate data yields the same number of extrema as originally observed. This procedure could only be applied to the cortex, which is topologically comparable to a sphere. The choice of an appropriate null model for subcortical areas is more complicated. While model-based procedures (for example, ref. 110) can generate surrogates with comparable spatial autocorrelation, they require parameter tuning, can show variable fits across individuals and do not preserve the exact values of the original data, making thresholding challenging. We therefore randomly shuffled deviation values across subcortical areas before thresholding them. Note that this null model is more lenient than a spatially constrained test since the model preserves fewer features of the data. As we found no significant differences using this test, our results will not change under a more stringent, spatially constrained null model.

We then obtained individual-specific surrogate FC union maps using the same procedure described above (Extended Data Fig. 3h), and calculated surrogate within-group overlap maps (Extended Data Fig. 3i) and between-group overlap difference maps (Extended Data Fig. 3j). This procedure was repeated 10,000 times to generate a null distribution of overlap difference maps for each disorder, again using the generalized Pareto tail approximation (Extended Data Fig. 3k). Statistically significant differences were identified using a threshold of $P_{\text{FDR}} < 0.05$, two tailed.

To summarize, group differences in circuit-level overlap were assessed with respect to two null models, one based on permutation of group labels and one based on spatial permutation of individual deviation maps. The group-based permutation test identifies regions showing differences in circuit-level overlap regardless of deviation burden. The spatial permutation test can be used to assess the degree to which such differences are driven by variations in deviation burden. More specifically, if a given brain region shows a significant difference for both tests, we have evidence to indicate that extreme deviations within the disorder preferentially aggregate within the functional circuitry of that region. If, on the other hand, a region shows a difference with the group-based permutation test but not the spatial permutation test, then the observed group differences are consistent with expectations from randomly selecting of the same number of seeds in each group, indicating that the effects observed under group-based permutation can be attributed to the fact that one group has a higher deviation burden than the other. Note that this scenario still yields important information about the extent of circuit-level heterogeneity in the disorder, since we generally expect cases to have a higher deviation burden and any resulting circuit-level overlap will still have phenotypic consequences regardless of whether such overlaps are explained by deviation burden or not. The spatial permutation test simply offers additional insights into the potential mechanisms that may drive group differences in circuit-level overlap.

Evaluating heterogeneity across spatial scales

To examine heterogeneity across spatial scales, we subtracted the regional Δ overlap map (Fig. 2g) from the circuit-level Δ overlap map (Fig. 3f) for each clinical group separately and evaluated the magnitude of the resulting differences using the group-based permutation testing. The Δ overlap maps quantify the difference in overlap observed at each region between each diagnostic group and controls. Thus, by subtracting the two Δ overlap maps, we directly test whether case-control differences in overlap are greater at the circuit compared with the regional level.

Mapping neural heterogeneity at the network level

The above procedure offers a means for understanding heterogeneity at the level of neural circuits that show strong FC with a deviant region. However, it is still possible for some pairs of brain regions to be affiliated with the same extended functional network despite being weakly coupled themselves (Fig. 1, right). We therefore sought to characterize neural heterogeneity at the level of the broader functional networks within which a given deviant region may be embedded.

Assigning regions to networks

We assigned each cortical region to one of seven canonical functional cortical networks using a well-validated network parcellation^{27,31} and assigned each subcortical region to either the medial temporal lobe (amygdala and hippocampus), thalamus or basal ganglia (nucleus accumbens, globus pallidus, putamen and caudate nucleus), as done previously²⁸, resulting in a total of ten distinct functional networks (Fig. 4a–c). The subcortical regions were assigned to coarse anatomical areas that were not aligned with the finely mapped cortical function networks because the correspondence between these subcortical nuclei, as parcellated here, and canonical cortical functional networks has not been extensively investigated. An alternative network assignment may influence network-level subcortical overlap values.

Characterizing network-level heterogeneity

To characterize network-level heterogeneity, we examined the degree to which GMV deviations aggregated within each network. For each network, we estimated the proportion of individuals within each diagnostic group that showed at least one extreme deviation in a region assigned to that network, separately for positive and negative deviations (Fig. 4d). We then computed case-control differences in overlap proportions for each network (Fig. 4e).

As in the circuit-level analysis, these observed group differences in network-level overlap were evaluated with respect to two complementary null models. The first involved a group-based randomization test in which a null distribution was generated by permuting group labels 10,000 times (Extended Data Fig. 3a–d). The second null model used the 10,000 spatially rotated maps generated in the circuit-level analysis to obtain surrogate estimates of group differences in network-level overlap while preserving variation in deviation burden (Extended Data Fig. 3e–g, l–o). In both cases, the tail of the null distribution was approximated using a generalized Pareto distribution and we used a significance threshold of $P_{\text{FDR}} < 0.05$, two tailed. As in the circuit-level analysis, these two null models allowed us to distinguish preferential network involvement from the effects of total deviation burden. Note that comparisons in overlap values between networks are complicated by differences in network size. Our analysis therefore focuses primarily on comparisons between cases and controls within networks since the control data provide a critical normative benchmark against which cases can be compared. We repeated the same analyses using a 20-network parcellation (17 cortical networks and 3 subcortical regions).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-023-01404-6>.

References

- Opel, N. et al. Cross-disorder analysis of brain structural abnormalities in six major psychiatric disorders: a secondary analysis of mega- and meta-analytical findings from the ENIGMA Consortium. *Biol. Psychiatry* **88**, 678–686 (2020).
- Hibar, D. P. et al. Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. *Mol. Psychiatry* **23**, 932–942 (2018).
- Boedhoe, P. S. W. et al. An empirical comparison of meta- and mega-analysis with data from the ENIGMA Obsessive-Compulsive Disorder Working Group. *Front. Neuroinform.* **12**, 1–8 (2019).
- Hoogman, M. et al. Brain imaging of the cortex in ADHD: a coordinated analysis of large-scale clinical and population-based samples. *Am. J. Psychiatry* **176**, 531–542 (2019).
- van Erp, T. G. M. et al. Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol. Psychiatry* **84**, 644–654 (2018).
- Schmaal, L. et al. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Mol. Psychiatry* **22**, 900–909 (2017).
- Van Rooij, D. et al. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: results from the ENIGMA ASD Working Group. *Am. J. Psychiatry* **175**, 359–369 (2018).
- Fried, E. I. & Nesse, R. M. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. *J. Affect. Disord.* **172**, 96–102 (2015).
- Young, G., Lareau, C. & Pierre, B. One quintillion ways to have PTSD comorbidity: recommendations for the disordered DSM-5. *Psychol. Inj. Law* **7**, 61–74 (2014).
- Wolfers, T. et al. Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry* **75**, 1146–1155 (2018).
- Marquand, A. F., Rezek, I., Buitelaar, J. & Beckmann, C. F. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biol. Psychiatry* **80**, 552–561 (2016).
- Rutherford, S. et al. The normative modeling framework for computational psychiatry. *Nat. Protoc.* **17**, 1711–1734 (2022).
- Wolfers, T. et al. Individual differences v. the average patient: mapping the heterogeneity in ADHD using normative models. *Psychol. Med.* **50**, 314–323 (2019).
- Wolfers, T. et al. Replicating extensive brain structural heterogeneity in individuals with schizophrenia and bipolar disorder. *Hum. Brain Mapp.* **42**, 2546–2555 (2021).
- Lv, J. et al. Individual deviations from normative models of brain structure in a large cross-sectional schizophrenia cohort. *Mol. Psychiatry* **26**, 3512–3523 (2021).
- Floris, D. L. et al. Atypical brain asymmetry in autism—a candidate for clinically meaningful stratification. *Biol. Psychiatry. Cogn. Neurosci. Neuroimaging* **6**, 802–812 (2021).
- Zabihi, M. et al. Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **4**, 567–578 (2019).
- Bayer, J. M. M. et al. Accommodating site variation in neuroimaging data using normative and hierarchical Bayesian models. *NeuroImage* <https://doi.org/10.1016/j.neuroimage.2022.119699> (2021).
- Carrera, E. & Tononi, G. Diaschisis: past, present, future. *Brain* **137**, 2408–2422 (2014).
- Fornito, A., Zalesky, A. & Breakspear, M. The connectomics of brain disorders. *Nat. Rev. Neurosci.* **16**, 159–172 (2015).
- von Monakow, C. *Die Lokalisation im Grosshirn und der Abbau der Funktion durch kortikale Herde* (JF Bergmann, 1914).
- Boes, A. D. et al. Network localization of neurological symptoms from focal brain lesions. *Brain* **138**, 3061–3075 (2015).
- Fox, M. D. Mapping symptoms to brain networks with the human connectome. *N. Engl. J. Med.* **379**, 2237–2245 (2018).
- Joutsa, J., Corp, D. T. & Fox, M. D. Lesion network mapping for symptom localization: recent developments and future directions. *Curr. Opin. Neurol.* **35**, 453–459 (2022).
- Gaser, C. & Dahnke, R. CAT—a computational anatomy toolbox for the analysis of structural MRI data. *Hbm* **2016**, 336–348 (2016).
- Ashburner, J. & Friston, K. J. Voxel-based morphometry—the methods. *NeuroImage* **11**, 805–821 (2000).
- Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
- Tian, Y., Margulies, D. S., Breakspear, M. & Zalesky, A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nat. Neurosci.* **23**, 1421–1432 (2020).
- Kia, S. M. et al. Hierarchical Bayesian regression for multi-site normative modeling of neuroimaging data. *Lect. Notes Comput. Sci.* **12267 LNCS**, 699–709 (2020).
- Seghier, M. L. & Price, C. J. Visualising inter-subject variability in fMRI using threshold-weighted overlap maps. *Sci. Rep.* **6**, 1–13 (2016).
- Yeo, B. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
- Insel, T. R. Faulty circuits. *Sci. Am.* **302**, 44–52 (2010).
- Cortese, S. et al. Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI studies. *Am. J. Psychiatry* **169**, 1038–1055 (2012).
- Fornito, A., Zalesky, A., Pantelis, C. & Bullmore, E. T. Schizophrenia, neuroimaging and connectomics. *NeuroImage* **62**, 2296–2314 (2012).
- Goodkind, M. et al. Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* **72**, 305–315 (2015).
- McTeague, L. M. et al. Identification of common neural circuit disruptions in cognitive control across psychiatric disorders. *Am. J. Psychiatry* **174**, 676–685 (2017).
- McTeague, L. M. et al. Identification of common neural circuit disruptions in emotional processing across psychiatric disorders. *Am. J. Psychiatry* **177**, 411–421 (2020).
- Ham, T., Leff, A., de Boissezon, X., Joffe, A. & Sharp, D. J. Cognitive control and the salience network: an investigation of error processing and effective connectivity. *J. Neurosci.* **33**, 7091–7098 (2013).
- Chong, J. S. X., Ng, G. J. P., Lee, S. C. & Zhou, J. Salience network connectivity in the insula is associated with individual differences in interoceptive accuracy. *Brain Struct. Funct.* **222**, 1635–1644 (2017).
- Menon, V. & Uddin, L. Q. Salience, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).

41. Uddin, L. Q. et al. Salience network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiatry* **70**, 869–879 (2013).
42. Gürsel, D. A., Avram, M., Sorg, C., Brandl, F. & Koch, K. Frontoparietal areas link impairments of large-scale intrinsic brain networks with aberrant fronto-striatal interactions in OCD: a meta-analysis of resting-state functional connectivity. *Neurosci. Biobehav. Rev.* **87**, 151–160 (2018).
43. Koch, S. B. J. et al. Aberrant resting-state brain activity in posttraumatic stress disorder: a meta-analysis and systematic review. *Depress. Anxiety* **33**, 592–605 (2016).
44. Wang, X., Zhang, W., Sun, Y., Hu, M. & Chen, A. Aberrant intra-salience network dynamic functional connectivity impairs large-scale network interactions in schizophrenia. *Neuropsychologia* **93**, 262–270 (2016).
45. Shanmugan, S. et al. Common and dissociable mechanisms of executive system dysfunction across psychiatric disorders in youth. *Am. J. Psychiatry* **173**, 517–526 (2016).
46. Caspi, A. et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin. Psychol. Sci.* **2**, 119–137 (2014).
47. Ahmari, S. E. & Rauch, S. L. The prefrontal cortex and OCD. *Neuropsychopharmacology* <https://doi.org/10.1038/s41386-021-01130-2> (2021).
48. Pizzagalli, D. A. & Roberts, A. C. Prefrontal cortex and depression. *Neuropsychopharmacology* <https://doi.org/10.1038/s41386-021-01101-7> (2021).
49. Zhou, Y., Fan, L., Qiu, C. & Jiang, T. Prefrontal cortex and the dysconnectivity hypothesis of schizophrenia. *Neurosci. Bull.* **31**, 207–219 (2015).
50. Kenwood, M. M., Kalin, N. H. & Barbas, H. The prefrontal cortex, pathological anxiety, and anxiety disorders. *Neuropsychopharmacology* <https://doi.org/10.1038/s41386-021-01109-z> (2021).
51. Koenigs, M. & Grafman, J. The functional neuroanatomy of depression: distinct roles for ventromedial and dorsolateral prefrontal cortex. *Behav. Brain Res.* **201**, 239–243 (2009).
52. O'Reardon, J. P. et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol. Psychiatry* **62**, 1208–1216 (2007).
53. Nguyen, T. D., Hieronymus, F., Lorentzen, R., McGirr, A. & Østergaard, S. D. The efficacy of repetitive transcranial magnetic stimulation (rTMS) for bipolar depression: a systematic review and meta-analysis. *J. Affect. Disord.* **279**, 250–255 (2021).
54. Padmanabhan, J. L. et al. A human depression circuit derived from focal brain lesions. *Biol. Psychiatry* **86**, 749–758 (2019).
55. Siddiqi, S. H. et al. Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. *Nat. Hum. Behav.* **5**, 1707–1716 (2021).
56. Siddiqi, S. H., Kording, K. P., Parvizi, J. & Fox, M. D. Causal mapping of human brain function. *Nat. Rev. Neurosci.* **23**, 361–375 (2022).
57. Buckner, R. L. et al. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J. Neurosci.* **29**, 1860–1873 (2009).
58. van den Heuvel, M. P. & Sporns, O. Network hubs in the human brain. *Trends Cogn. Sci.* **17**, 683–696 (2013).
59. Arnatkeviciute, A. et al. Genetic influences on hub connectivity of the human connectome. *Nat. Commun.* **12**, 4237 (2021).
60. Schumann, C. M. et al. Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *J. Neurosci.* **30**, 4419–4427 (2010).
61. Palomero-Gallagher, N. et al. Human pregenual anterior cingulate cortex: structural, functional, and connectional heterogeneity. *Cereb. Cortex* **29**, 2552–2574 (2019).
62. Miller, E. K. The prefrontal cortex: categories, concepts, and cognitive control. *Nat. Rev. Neurosci.* **1**, 59–65 (2000).
63. Chopra, S. et al. Differentiating the effect of antipsychotic medication and illness on brain volume reductions in first-episode psychosis: a longitudinal, randomised, triple-blind, placebo-controlled MRI study. *Neuropsychopharmacology* **46**, 1494–1501 (2021).
64. Winter, N. R. et al. Quantifying deviations of brain structure and function in major depressive disorder across neuroimaging modalities. *JAMA Psychiatry* **79**, 879 (2022).
65. Rutherford, S. et al. Charting brain growth and aging at high spatial precision. *elife* **11**, e72904 (2022).
66. Bethlehem, R. A. I. et al. Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).
67. Tiego, J. et al. Precision behavioral phenotyping as a strategy for uncovering the biological correlates of psychopathology. *Nat. Ment. Health* **1**, 304–315 (2023).
68. Kraus, B. et al. Insights from personalized models of brain and behavior for identifying biomarkers in psychiatry. *Neurosci. Biobehav. Rev.* **152**, 105259 (2023).
69. Kotov, R. et al. The hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* **126**, 454–477 (2017).
70. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth edition* (American Psychiatric Association, 2013).
71. The ICD-10 Classification of Mental and Behavioural Disorders (World Health Organization, Geneva, 1992).
72. Bellon, E. M. et al. MR artifacts: a review. *Am. J. Roentgenol.* **147**, 1271–1281 (1986).
73. Zaitsev, M., Maclaren, J. & Herbst, M. Motion artifacts in MRI: a complex problem with many partial solutions. *J. Magn. Reson. Imaging* **42**, 887–901 (2015).
74. Gaser, C., Dahnke, R., Thompson, P. M., Kurth, F., & Luders, E. CAT—a computational anatomy toolbox for the analysis of structural MRI data. *BioRxiv* <https://doi.org/10.1101/2022.06.11.495736> (2022).
75. SPM12. *Univ. College London* <http://www.fil.ion.ucl.ac.uk/spm/software/spm12> (2020).
76. The MathWorks Inc. MATLAB version 9.4.0.813654 (R2018a). (The MathWorks Inc., 2018).
77. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113 (2007).
78. Kong, R. et al. Individual-specific areal-level parcellations improve functional connectivity prediction of behavior. *Cereb. Cortex* **31**, 4477–4500 (2021).
79. get_totals.m. *Univ. College London* http://www0.cs.ucl.ac.uk/staff/gridway/vbm/get_totals.m
80. amarquand. *GitHub* <https://github.com/amarquand/PCNtoolkit.git>
81. Focke, N. K. et al. Multi-site voxel-based morphometry—not quite there yet. *NeuroImage* **56**, 1164–1170 (2011).
82. Box, G. E. P. & Cox, D. R. An analysis of transformations. *J. R. Stat. Soc. B* **26**, 211–243 (1964).
83. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
84. Winkler, A. M., Ridgway, G. R., Douaud, G., Nichols, T. E. & Smith, S. M. Faster permutation inference in brain imaging. *NeuroImage* **141**, 502–516 (2016).
85. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *NeuroImage* **92**, 381–397 (2014).
86. Vanasse, T. J. et al. Brain pathology recapitulates physiology: a network meta-analysis. *Commun. Biol.* **4**, 301 (2021).

87. Van Essen, D. C. et al. The WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013).
88. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* **17**, 825–841 (2002).
89. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).
90. Andersson, J. L. R., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* **20**, 870–888 (2003).
91. Griffanti, L. et al. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage* **95**, 232–247 (2014).
92. Aquino, K. M., Fulcher, B. D., Parkes, L., Sabaroedin, K. & Fornito, A. Identifying and removing widespread signal deflections from fMRI data: rethinking the global signal regression problem. *NeuroImage* **212**, 116614–116614 (2020).
93. Burgess, G. C. et al. Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the Human Connectome Project. *Brain Connect.* **6**, 669–680 (2016).
94. Glasser, M. F. et al. Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. *NeuroImage* **181**, 692–717 (2018).
95. Chang, C. & Glover, G. H. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *NeuroImage* **47**, 1448–1459 (2009).
96. Saad, Z. S. et al. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain Connect.* **2**, 25–32 (2012).
97. Parkes, L., Fulcher, B., Yücel, M. & Fornito, A. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage* **171**, 415–436 (2018).
98. Power, J. D. et al. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341 (2014).
99. Matsui, T., Murakami, T. & Ohki, K. Transient neuronal coactivations embedded in globally propagating waves underlie resting-state functional connectivity. *Proc. Natl Acad. Sci. USA* **113**, 6556–6561 (2016).
100. Li, J. et al. Global signal regression strengthens association between resting-state functional connectivity and behavior. *NeuroImage* **196**, 126–141 (2019).
101. Fox, M. D., Zhang, D., Snyder, A. Z. & Raichle, M. E. The global signal and observed anticorrelated resting state brain networks. *J. Neurophysiol.* **101**, 3270–3283 (2009).
102. Silver, N. C. & Dunlap, W. P. Averaging correlation coefficients: should Fisher's z transformation be used? *J. Appl. Psychol.* **72**, 146–148 (1987).
103. Smith, S. M. & Nichols, T. E. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* **44**, 83–98 (2009).
104. Missing parcel in Schaefer 1000 conte69. *GitHub* <https://github.com/ThomasYeoLab/CBIG/issues/10> (2020).
105. Kuceyeski, A., Maruta, J., Relkin, N. & Raj, A. The network modification (NeMo) tool: elucidating the effect of white matter integrity changes on cortical and subcortical structural connectivity. *Brain Connect.* **3**, 451–463 (2013).
106. Sperber, C., Griffis, J. & Kasties, V. Indirect structural disconnection-symptom mapping. *Brain Struct. Funct.* **227**, 3129–3144 (2022).
107. Alexander-Bloch, A., Giedd, J. N. & Bullmore, E. Imaging structural co-variance between human brain regions. *Nat. Rev. Neurosci.* **14**, 322–336 (2013).
108. Kuhn, H. W. The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* **2**, 83–97 (1955).
109. Markello, R. D. & Misic, B. Comparing spatial null models for brain maps. *NeuroImage* **236**, 118052–118052 (2021).
110. Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. *NeuroImage* **220**, 117038–117038 (2020).

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

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

© The Author(s) 2023

¹Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Melbourne, Victoria, Australia. ²Monash Biomedical Imaging, Monash University, Melbourne, Victoria, Australia. ³Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA. ⁴Department of Psychiatry, Rutgers University, Piscataway, NJ, USA. ⁵School of Physics, University of Sydney, Sydney, New South Wales, Australia. ⁶BrainKey Inc, Palo alto, CA, USA. ⁷Donders Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, the Netherlands. ⁸Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands. ⁹Department of Psychiatry, University Medical Center Utrecht, Utrecht, the Netherlands. ¹⁰Department of Cognitive Science and Artificial Intelligence, Tilburg University, Tilburg, the Netherlands. ¹¹Norwegian Centre for Mental Disorders Research, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway. ¹²Department of Psychiatry and Psychotherapy, Tübingen Center for Mental Health (TÜCMH), University of Tübingen, Tübingen, Germany. ¹³Department of Psychiatry, Donders Institute of Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands. ¹⁴Department of Human Genetics, Donders Institute of Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands. ¹⁵Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, the Netherlands. ¹⁶Centre for Functional MRI of the Brain, Nuffield Department of Clinical Neurosciences, Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK. ¹⁷Department of Psychology, University of Oslo, Oslo, Norway. ¹⁸KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo and Oslo University Hospital, Oslo, Norway. ¹⁹Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne and Melbourne Health, Melbourne, Victoria, Australia. ²⁰Department of Biomedical Engineering, The University of Melbourne, Melbourne, Victoria, Australia. ²¹Department of Psychiatry, University of Melbourne, Melbourne, Victoria, Australia. ²²Department of Psychiatry, Bellvitge University Hospital, Bellvitge Biomedical Research Institute, Barcelona, Spain. ²³Centro de Investigación Biomédica en Red de Salud Mental, Carlos III Health Institute, Madrid, Spain. ²⁴Department of Social Psychology and

Quantitative Psychology, Universitat de Barcelona, Barcelona, Spain. ²⁵Sant Pau Mental Health Research Group, Institut d'Investigació Biomèdica Sant Pau, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain. ²⁶Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain. ²⁷QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ²⁸Australian Characterisation Commons at Scale (ACCS) Project, Monash eResearch Centre, Melbourne, Victoria, Australia. ²⁹Institute for Mental and Physical Health and Clinical Translation School of Medicine, Deakin University, Geelong, Victoria, Australia. ³⁰Orygen, Melbourne, Victoria, Australia. ³¹Centre for Youth Mental Health, University of Melbourne, Melbourne, Victoria, Australia. ³²Florey Institute for Neuroscience and Mental Health, Parkville, Victoria, Australia. ³³Department of Neuroimaging, Centre of Neuroimaging Sciences, Institute of Psychiatry, King's College London, London, UK. ³⁴These authors contributed equally: Andre F. Marquand, Alex Fornito. ✉e-mail: ashlea.segal@gmail.com; alex.fornito@monash.edu

Methods

Original study protocols were approved by the local ethics committee of each dataset, and written informed consent was obtained from each participant. This current study was approved by the Monash University Research Ethics Committee (project ID: 23534). In this study, data collection and analysis were not blind to the conditions of the experiment because participants were not placed into experimental groups.

Participants

This study included 3,746 individuals (1,865 HCs; 1,833 cases across six different diagnostic categories) from 14 separate, independently acquired datasets and 25 scan sites. No statistical methods were used to predetermine sample sizes but to our knowledge, our dataset is one of the largest aggregations of case–control data outside the ENIGMA consortium¹¹¹. Full details on the study design and clinical characteristics have been described previously for each dataset (for basic characteristics and relevant references, see Supplementary Table 1). Each study was approved by the relevant ethics committee and written informed consent was obtained from each participant. Participants were reimbursed for their time in accordance with the local research protocols and ethics committee guidelines.

The final sample was drawn from a larger pool of individuals recruited across the 14 datasets. In addition to the specific quality control procedures used within each study (references in Table 1), we performed a series of quality control checks and exclusions for our analysis. Specifically, we excluded participants if they were below 18 years or above 64 years of age ($N = 346$); did not have the necessary clinical data (clinical diagnosis or, for HCs, absence of any clinical diagnosis) or demographic information (age, sex and scanner site, $N = 53$); if their T1-weighted structural MRI scan did not survive our stringent manual and automated quality control procedure, as explained below ($N = 269$); or if the data came from a site with less than ten individuals in the same group and sex (described in the Normative model section, below; $N = 217$). Our final sample available for analysis thus comprised 1,465 HCs and 1,294 cases. Demographic and other details of this cohort are presented in Supplementary Table 1.

Data availability

Summary of data availability for each dataset used in this study is described below. Autism Brain Imaging Data Exchange I (ABIDE I)¹¹² and ABIDE II¹¹³ datasets are available through the ABIDE repository, http://fcon_1000.projects.nitrc.org/indi/abide/. The ASRB¹¹⁴ dataset is available through the ASRB repository, subject to approval of the ASRB Access Committee (<https://www.neura.edu.au/discovery-portal/asrb/>). First Episode Mania Study¹¹⁵, Monash Cohort¹¹⁶, Obsessive–Compulsive and Problematic Gambling study¹¹⁷, SPAINOCD¹¹⁸ and YoDA¹¹⁹ datasets are available from the principal investigators of the respective studies, subject to evaluation of the request and local ethics committee requirements. The International Multicentre persistent ADHD Collaboration (IMpACT-NL)¹²⁰ and TOP15¹²¹ datasets are not publicly available due to privacy or ethical restrictions. Kansas Musical Depression Study (KAN-MDD; ds000171)^{122,123}, Massachusetts Institute of Technology Autism Study (MITASD; ds000212.v1.0.0)^{124,125}, Russia fMRI Depression Study (RUSMDD; ds002748.v1.0.0)^{126,127} and University of Washington ASD Study (WASHASD; ds002522.v1.0.0)^{128,129} datasets are available through the OpenNeuro repository, <https://doi.org/10.18112/openneuro>. The HCP dataset is available in the Human Connectome Project repository (<https://www.humanconnectome.org/study/hcp-young-adult>).

Code availability

The PCN toolkit (version 0.16, <https://github.com/amarquand/PCN-toolkit.git>) was used to generate the person-specific z-score deviation maps. The HCP Pipelines: Seed fMRI Analysis (<https://github.com/Washington-University/HCPpipelines.git>) was adapted to map the regions functionally coupled to extreme deviations. Figures were

generated programmatically¹³⁰, with cortical surface renderings generated using plotSurfaceROIboundary (<https://github.com/StuartJO/plotSurfaceROIboundary.git>). All code used to produce the figures for this study can be found at https://github.com/ashlea-segal/crossdis-order_multiscale_gmv_heterogeneity.git.

References

- Thompson, P. M. et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav.* **8**, 153–182 (2014).
- Di Martino, A. et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* **19**, 659–667 (2014).
- Di Martino, A. et al. Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci. Data* **4**, 1–15 (2017).
- Loughland, C. et al. Australian Schizophrenia Research Bank: a database of comprehensive clinical, endophenotypic and genetic data for aetiological studies of schizophrenia. *Aust. N.Z. J. Psychiatry* **44**, 1029–1035 (2010).
- Dandash, O. et al. Differential effect of quetiapine and lithium on functional connectivity of the striatum in first episode mania. *Transl. Psychiatry* **8**, 59 (2018).
- Sabaroedin, K. et al. Functional connectivity of corticostriatal circuitry and psychosis-like experiences in the general community. *Biol. Psychiatry* **86**, 16–24 (2019).
- Parkes, L. et al. Transdiagnostic variations in impulsivity and compulsivity in obsessive-compulsive disorder and gambling disorder correlate with effective connectivity in cortical-striatal-thalamic-cortical circuits. *NeuroImage* **202**, 116070–116070 (2019).
- Real, E. et al. Brain structural correlates of obsessive–compulsive disorder with and without preceding stressful life events. *World J. Biol. Psychiatry* **17**, 366–377 (2016).
- Davey, C. G., Cearnas, M., Jamieson, A. & Harrison, B. J. Suppressed activity of the rostral anterior cingulate cortex as a biomarker for depression remission. *Psychol. Med.* <https://doi.org/10.1017/S0033291721004323> (2021).
- Hoogman, M. et al. Nitric oxide synthase genotype modulation of impulsivity and ventral striatal activity in adult ADHD patients and healthy comparison subjects. *Am. J. Psychiatry* **168**, 1099–1106 (2011).
- Doan, N. T. et al. Distinct multivariate brain morphological patterns and their added predictive value with cognitive and polygenic risk scores in mental disorders. *NeuroImage Clin.* **15**, 719–731 (2017).
- Lepping, R. J. et al. Neural processing of emotional musical and nonmusical stimuli in depression. *OpenNeuro*. [Dataset] (2018).
- Lepping, R. J. et al. Neural processing of emotional musical and nonmusical stimuli in depression. *PLoS ONE* **11**, 1–23 (2016).
- Koster-Hale, J., Saxe, R., Dungan, J. & Young, L. L. Decoding moral judgments from neural representations of intentions. *Proc. Natl Acad. Sci. USA* **110**, 5648–5653 (2013).
- Young, L. et al. Moral judgments of intentional and accidental moral violations across Harm and Purity domains. *OpenNeuro*. [Dataset] <https://doi.org/10.18112/openneuro.ds000212.v1.0.0> (2019).
- Bezmaternykh, D., Melnikov, M., Savelov, A. & Petrovskii, E. Resting state with closed eyes for patients with depression and healthy participants. *OpenNeuro*. [Dataset] <https://doi.org/10.18112/openneuro.ds002748.v1.0.5> (2020).
- Mel'nikov, M. E. et al. fMRI response of parietal brain areas to sad facial stimuli in mild depression. *Bull. Exp. Biol. Med.* **165**, 741–745 (2018).

128. Kolodny, T., Schallmo, M.-P. & Murray, S. O. Contrast response functions. *OpenNeuro*. [Dataset] <https://doi.org/10.18112/openneuro.ds002522.v1.0.0> (2020).
129. Kolodny, T., Schallmo, M. P., Gerdts, J., Bernier, R. A. & Murray, S. O. Response dissociation in hierarchical cortical circuits: a unique feature of autism spectrum disorder. *J. Neurosci.* **40**, 2269–2281 (2020).
130. Chopra, S., Labache, L., Dhamala, E., Orchard, E. R. & Holmes, A. A practical guide for generating reproducible and programmatic neuroimaging visualizations. Preprint at *OSF Preprints* <https://doi.org/10.31219/osf.io/mtvy9> (2023).

Acknowledgements

This work was supported by the MASSIVE HCP facility (<http://www.massive.org.au>). We thank B. Fulcher for helpful discussions about this manuscript, S. Oldham and S. Chopra for sharing code to generate the cortical and subcortical brain surface renderings for the figures, and M.L. Seghier for sharing code to generate threshold-weighted overlap maps. This research was supported by the following grants: US Department of Health & Human Services | NIH | National Institute of Mental Health (NIMH) K99MH127296, L.P.; 2020 NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation, L.P.; German Research Foundation (DFG; project number(s): 513851350 and 390727645), T.W.; Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research) 016130669, B.F.; grant U54 EB020403 to the ENIGMA Consortium from the BD2K Initiative, a cross-NIH partnership, European College of Neuropsychopharmacology (ECNP) Network 'ADHD Across the Lifespan'. B.F.; Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research) 91619115, M.H.; Wellcome Trust (Wellcome) 215698/Z/19/Z, C.F.B.; Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research) Vici grant no. 17854 and NWO-CAS grant no. 012-200-013, C.F.B.; Norges Forskningsråd (Research Council of Norway) 249795, L.T.W.; Norges Forskningsråd (Research Council of Norway) 298646, L.T.W.; Norges Forskningsråd (Research Council of Norway) 300767, L.T.W.; South-Eastern Norway Regional Health Authority (2018076, 2019101), L.T.W.; European Research Council ('BRAINMINT' 802998), L.T.W.; Norges Forskningsråd (Research Council of Norway) 223273, O.A.A.; Norges Forskningsråd (Research Council of Norway) 276082, O.A.A.; Department of Health | NHMRC 1118153, A.Z.; Department of Health | NHMRC 1064643, B.J.H.; Department of Health | NHMRC 1124472, B.J.H.; Department of Health | NHMRC 1024570, C.G.D.; Department of Health | NHMRC 114738, C.G.D.; Generalitat de Catalunya (Government of Catalonia) SLT006/17/00249, C.S.; Carlos III Health Institute (PI16/00889 and PI19/01171), C.S.; Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (2021 SGR 01017), C.S.; Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (2021 SGR 00832), N.C.; Carlos III Health Institute (PI18/00036 and PI21/01756), N.C.; Turner Impact Fellowship from the Turner Institute for Brain and Mental Health, J.T.; Department of Health | NHMRC 1117188, M.Y.; Department of Health | NHMRC 1156072, M.B.; Department of Health | NHMRC 1136344, S.C.; Department of Health | NHMRC 1154378, M.A.B.; Department of Health | NHMRC 1146292, M.A.B.; Department of Health | NHMRC 1045354, M.A.B.; Department of Health | NHMRC 1006573, M.A.B.; European Research Council (ERC, grant 'MENTALPRECISION' 10100118), A.F.M.; Sylvia and Charles Viertel Charitable Foundation (Viertel Charitable Foundation), A.F.; Department of Health | NHMRC 1197431, A.F.; Department of Health | NHMRC 1146292, A.F.; Department of Health | NHMRC 1050504, A.F. We thank all the sites and investigators who have worked to share their data through data sharing repositories including ABIDE and OpenNeuro. This study was supported by the Australian Schizophrenia Research Bank (ASRB), which is supported by the National Health and

Medical Research Council of Australia, the Pratt Foundation, Ramsay Health Care, the Viertel Charitable Foundation and the Schizophrenia Research Institute. Additional data were provided (in part) by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Finally, we would like to thank Nathan Spreng, Michel Thiebaut de Schotten and the anonymous reviewer who provided valuable feedback.

Author contributions

A.S., A.F. and A.F.M. conceived the study and wrote the manuscript, with valuable revision from all authors. A.S. performed the formal analysis, with contributions from L.P., K.A. and S.M.K. Results were interpreted by A.S. and A.F., with contributions from A.F.M., L.P., S.M.K., A.Z. and J.T. The following provided data: T.W., B.F., M.H., C.F.B., L.T.W., O.A.A., A.Z., B.J.H., C.G.D., C.S.M., N.C., J.T., M.Y., L.B., C.S., M.B., S.C., M.A.B. and A.F. The project administrator was A.F.

Competing interests

K.A. is a scientific advisor to and shareholder in BrainKey Inc., a medical image analysis software company. B.F. has received educational speaking fees from Medice GmbH. C.F.B. is director and shareholder of SBGNeuro Ltd. O.A.A. is a consultant to HealthLytix and received speaker's honorarium from Lundbeck and Sunovion. N.C. participated in advisory boards and received speaker's honoraria from Angelini, Esteve, Janssen, Lundbeck, Novartis, Pfizer and Viatrix. Furthermore, they have been awarded research grants from the Ministry of Health, Ministry of Science and Innovation (CIBERSAM), and the Strategic Plan for Research and Innovation in Health (PERIS) for the period 2016–2020, as well as from Recercaixa and Marato TV3. M.Y. has received philanthropic donations from the David Winston Turner Endowment Fund, Wilson Foundation, as well as payments in relation to court, expert witness and/or expert review reports. Finally, he has received funding to conduct sponsored Investigator-Initiated trials (including Incannex Healthcare Ltd). These funding sources had no role in the design, management, data analysis, presentation or interpretation and write-up of the data. M.Y. also sits on the Advisory Boards of Centre of The Urban Mental Health, University of Amsterdam; Enosis Therapeutics; and Monash Biomedical Imaging Centre. M.B. has received grant/research support from the NIH, Cooperative Research Centre, Simons Autism Foundation, Cancer Council of Victoria, Stanley Medical Research Foundation, Medical Benefits Fund, National Health and Medical Research Council, Medical Research Futures Fund, Beyond Blue, Rotary Health, A2 milk company, Meat and Livestock Board, Woolworths, Avant and the Harry Windsor Foundation; has been a speaker for Abbot, AstraZeneca, Janssen and Janssen, Lundbeck and Merck; and served as a consultant to Allergan, AstraZeneca, Bioadvantex, Bionomics, Collaborative Medicinal Development, Eisai, Janssen and Janssen, Lundbeck Merck, Pfizer and Servier—all unrelated to this work. M.B. has received grant/research support from National Health and Medical Research Council, Wellcome Trust, Medical Research Future Fund, Victorian Medical Research Acceleration Fund, Centre for Research Excellence CRE, Victorian Government Department of Jobs, Precincts and Regions and Victorian COVID-19 Research Fund. He received honoraria from Springer, Oxford University Press, Cambridge University Press, Allen and Unwin, Lundbeck, Controversias Barcelona, Servier, Medisquire, HealthEd, ANZJP, EPA, Janssen, Medplan, Milken Institute, RANZCP, Abbott India, ASCP, Headspace and Sandoz. The other authors report no conflicts of interest.

Additional information

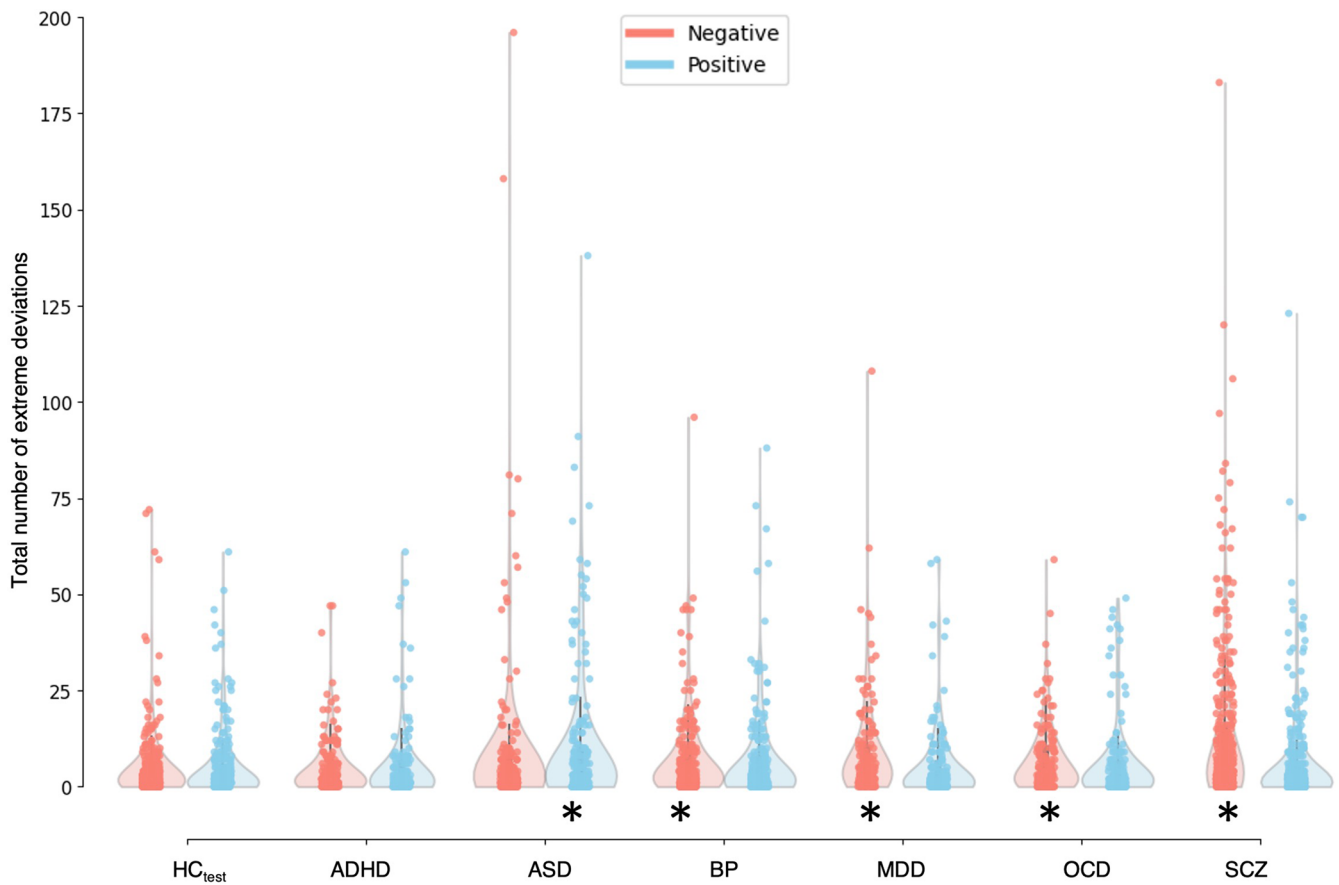
Extended data is available for this paper at <https://doi.org/10.1038/s41593-023-01404-6>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41593-023-01404-6>.

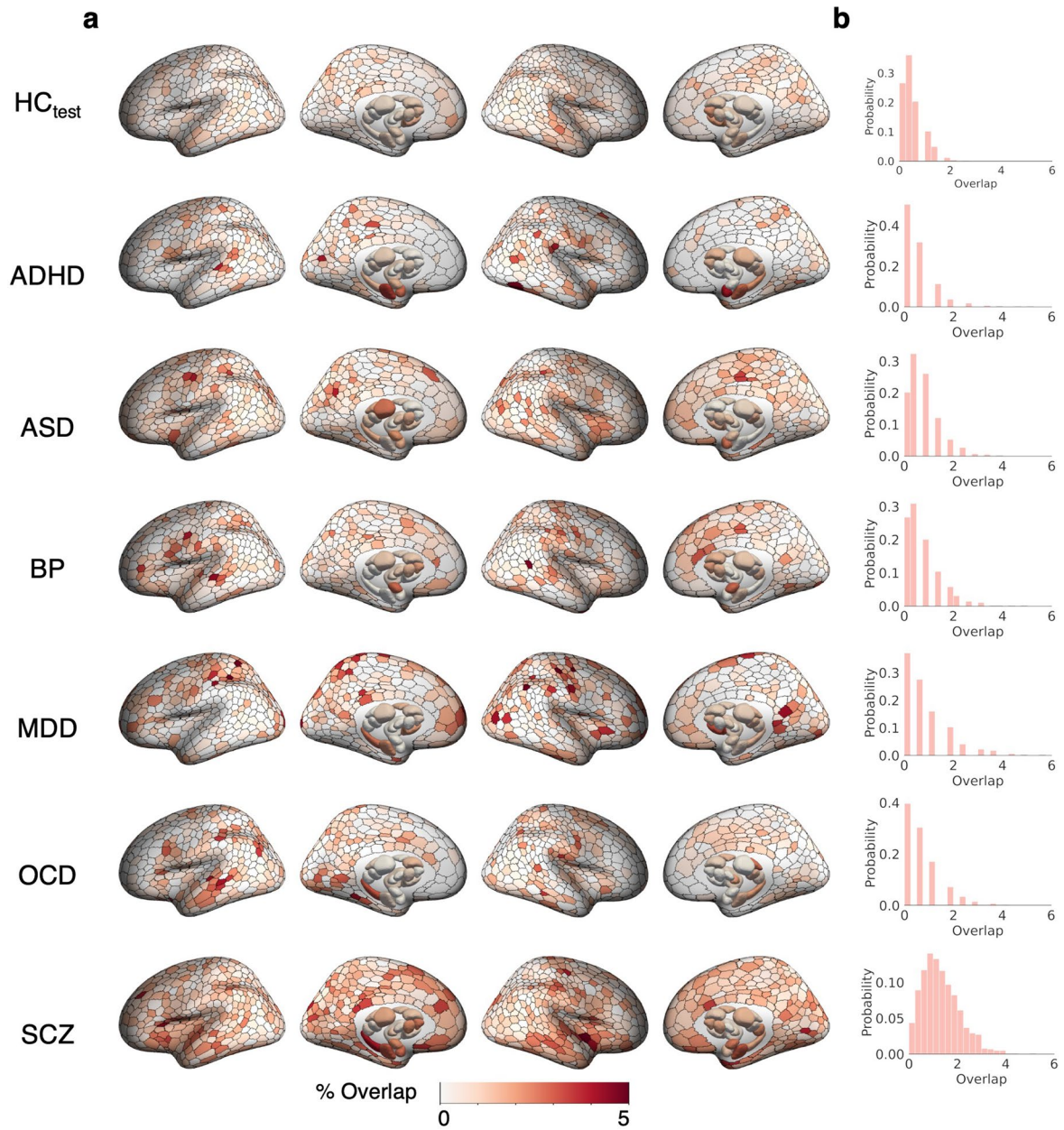
Correspondence and requests for materials should be addressed to Ashlea Segal or Alex Fornito.

Peer review information *Nature Neuroscience* thanks R. Nathan Spreng, Michel Thiebaut de Schotten, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

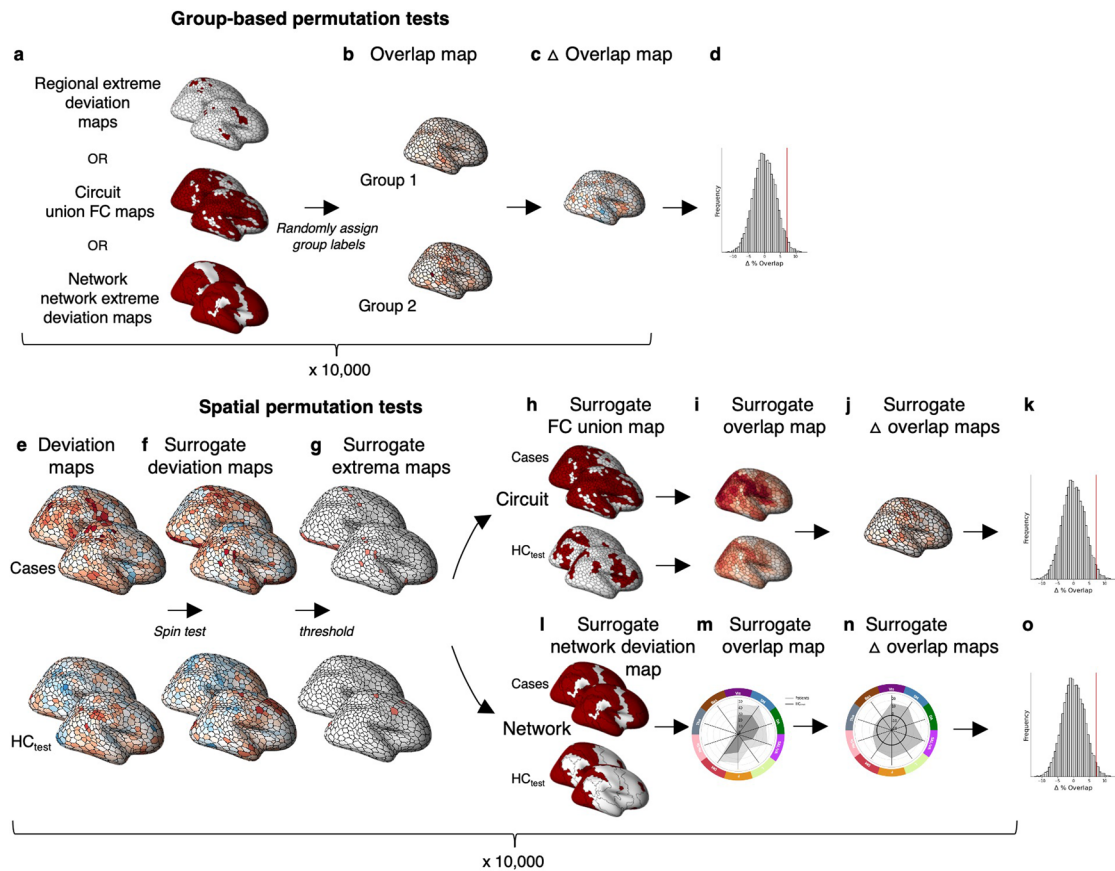
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | Distribution of person-specific positive ($Z > 2.6$; blue) and negative ($Z < -2.6$; red) deviation burden scores (that is, total number of extreme deviations) in each diagnostic group. * Indicates clinical groups showing a statistically significant difference in extreme deviation burden compared to the HC_{test} group (Mann-Whitney U-test, $p < .05$, two-tailed).

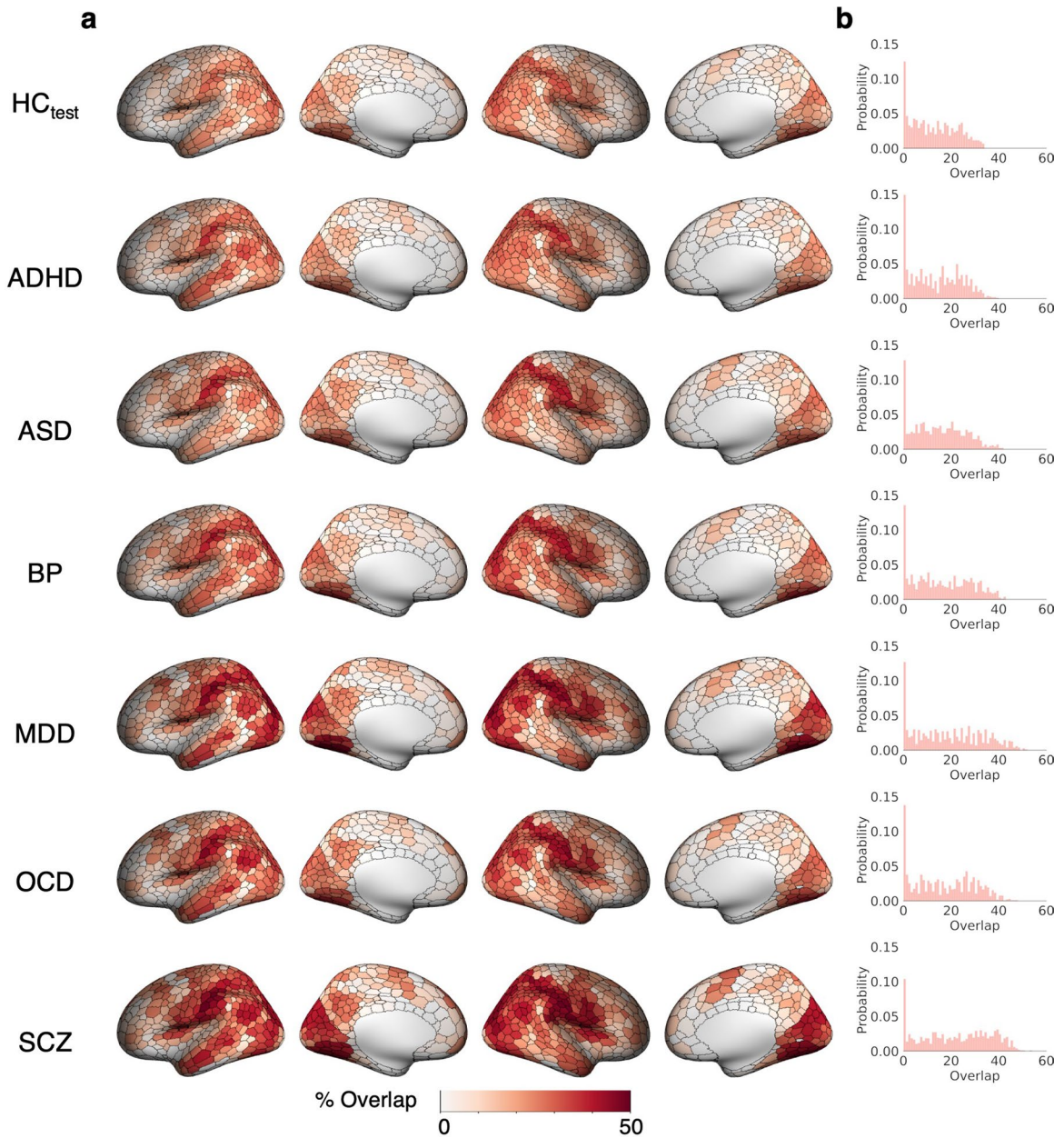


Extended Data Fig. 2 | Spatial overlap of extreme negative GMV deviations ($Z < -2.6$) in each group. a) Cortical and subcortical surface renderings showing spatial of overlap in 1032 brain regions, and b) the distribution of overlap percentages observed across all regions. Data used to generate this figure can be found in Supplementary Data 1 (*Regional_neg_thr26*).



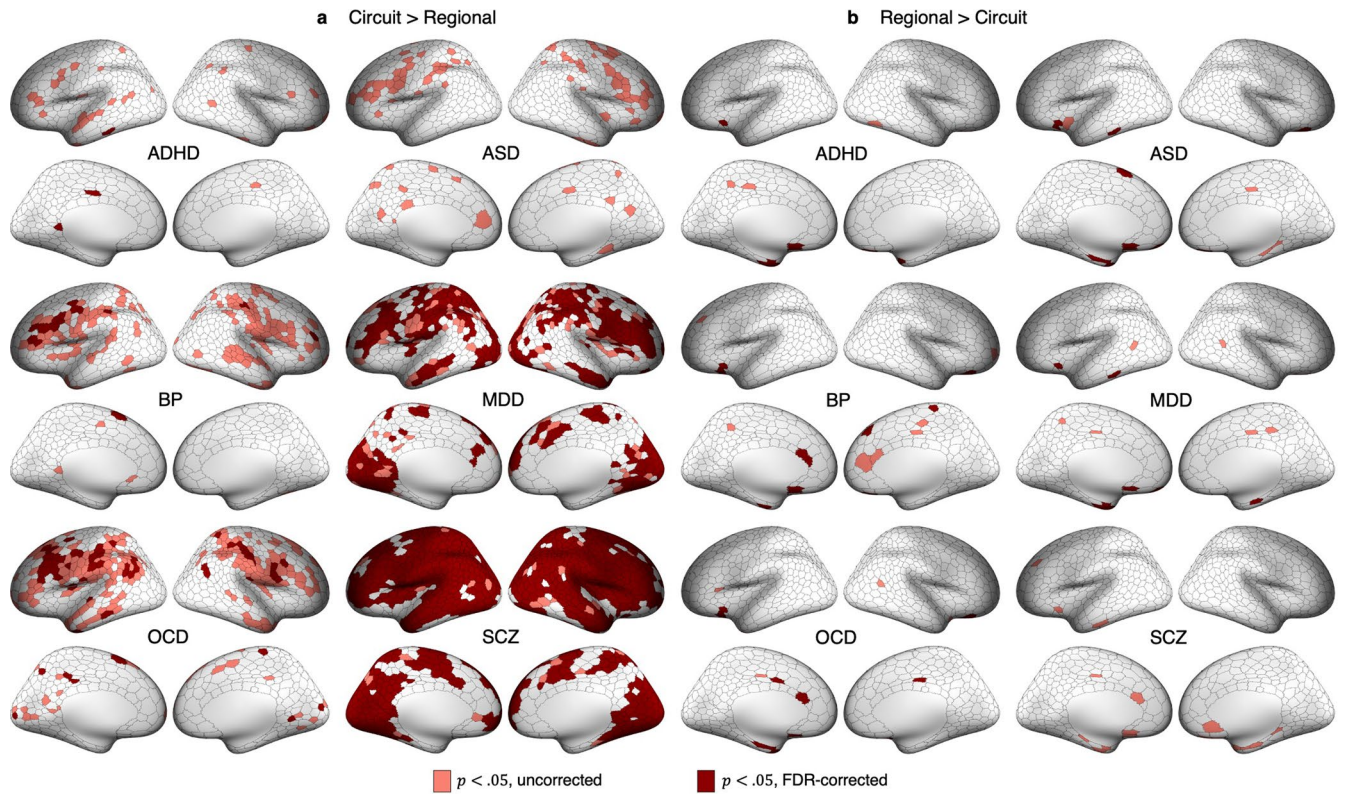
Extended Data Fig. 3 | Permutation tests for evaluating region-level, circuit-level, and network-level overlap. We used two types of permutation test to evaluate different hypotheses. Group-based permutation tests were used to evaluate group differences in region-level, circuit-level, and network-level overlap, regardless of total deviation burden. These tests involved repeating each analysis 10,000 times after shuffling case and control labels. (a) At each iteration, we obtained a new grouping of person-specific deviation maps according to the shuffled group labels. At the regional level, we focused on extreme deviation maps (Fig. 2e); at the circuit-level, we focused on union FC maps (Fig. 3d); and at the network level we focused on extreme network deviation maps (Fig. 4c). (b) For each brain region, we computed an overlap map for the HC_{test} and each clinical group under shuffled group assignment (*Overlap map*). (c) We then subtracted the surrogate HC_{test} overlap map from the surrogate clinical group's overlap map to obtain an overlap difference map (*Overlap map*). Steps (b) and (c) were repeated 10,000 times to derive an empirical distribution of overlap difference maps under the null hypothesis of random group assignment (d). For each brain region, we obtained *p*-values as the proportion of null values that exceeded the observed difference. The second type of permutation test we used was a spatial permutation test. (e) We used the unthresholded deviation maps of each person derived from the normative model to generate an ensemble of surrogate deviation maps for each individual in the test data (f). For cortical

regions, the surrogate maps were generated using Hungarian spin tests^{19,20}. For subcortical regions, we randomly shuffled deviation values across all subcortical areas (see Methods). (g) We then thresholded the null deviation maps ($Z > |2.6|$) to generate surrogate extreme deviation maps. To evaluate circuit-level group differences in overlap, (h) we obtained individual-specific surrogate FC union maps using the same procedure described in Fig. 3a-d. (i) For the HC_{test} group and each clinical group, we calculated surrogate within-group overlap maps. (j) We subtracted the HC_{test} surrogate FC overlap map from each clinical group's surrogate FC overlap map to obtain a surrogate overlap difference map (*Overlap map*). Steps (f) – (j) were repeated 10,000 times to generate (k) a null distribution of circuit-level overlap difference maps for each disorder. To evaluate network-level group differences in overlap, (l) we obtained surrogate network-level extreme deviation maps using the same procedure described in Figure 5a-c. (m) For each clinical group and the control group, we quantified the proportion of individuals showing a surrogate deviation within each network (*Overlap map*). (n) We subtracted the HC_{test} surrogate network overlap map from each clinical group's surrogate overlap (*Overlap map*). Steps (f) – (n) were repeated 10,000 times to generate (o) a null distribution of overlap difference maps for each disorder. For all tests, statistically significance differences were identified using a threshold of $p_{FDR} < 0.05$, two-tailed.



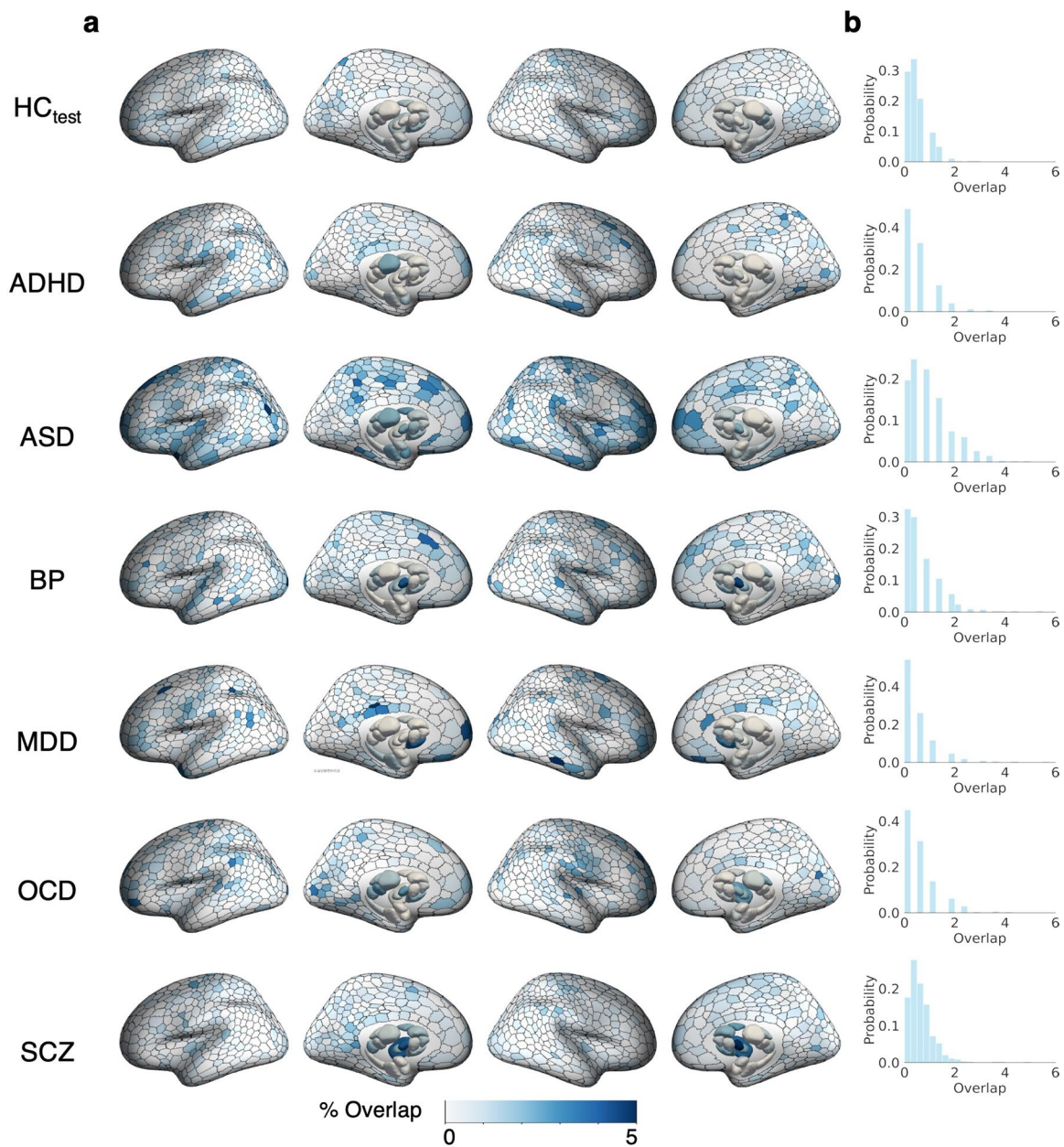
Extended Data Fig. 4 | Spatial overlap of regions functionally coupled (vertex-wise threshold $p_{FWE} < 0.025$), to extreme negative deviations ($Z < -2.6$) across groups, using a parcel-mapping threshold of 50%). a) Cortical

surface renderings showing spatial overlap and b) the distribution of overlap percentages observed across all regions. Data used to generate this figure can be found in Supplementary Data 1 (*Circuit_neg_parc50*).

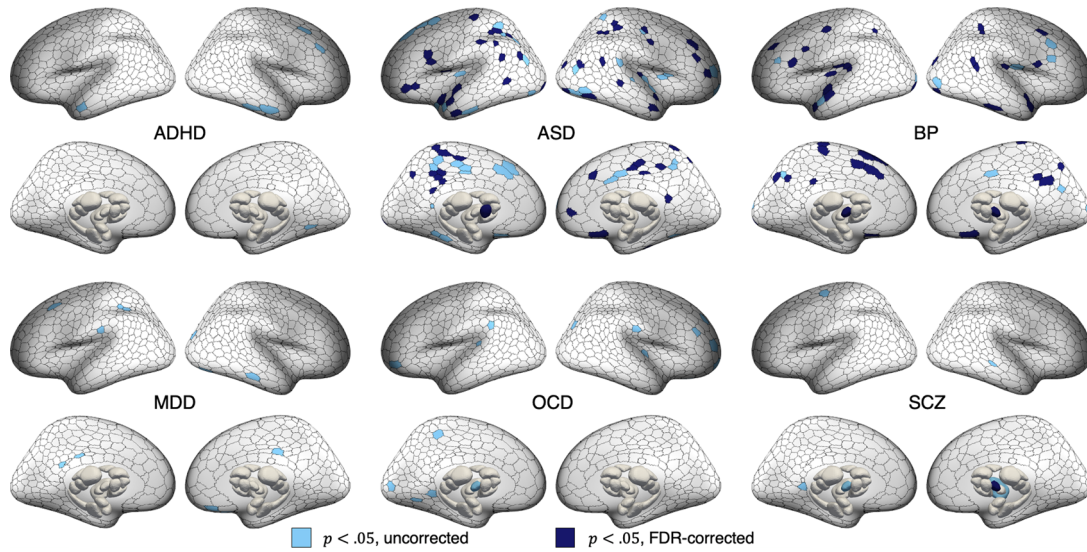


Extended Data Fig. 5 | Comparison of regional and circuit-level heterogeneity. (a) Statistical maps showing regions where the case-control differences in overlap observed at the circuit level (that is, Fig. 3g) is significantly greater than the overlap difference observed at the regional level (for example, Fig. 2h), identified using the group-based permutation testing (pink corresponds

to $p_{uncorrected} < .05$, red corresponds to $p_{FDR} < .05$, two-tailed, regional < circuit level overlap). (b) shows results for the reverse contrast (that is, regional > circuit level overlap). Data used to generate this figure can be found in Supplementary Data 1 (Regional vs Circuit).

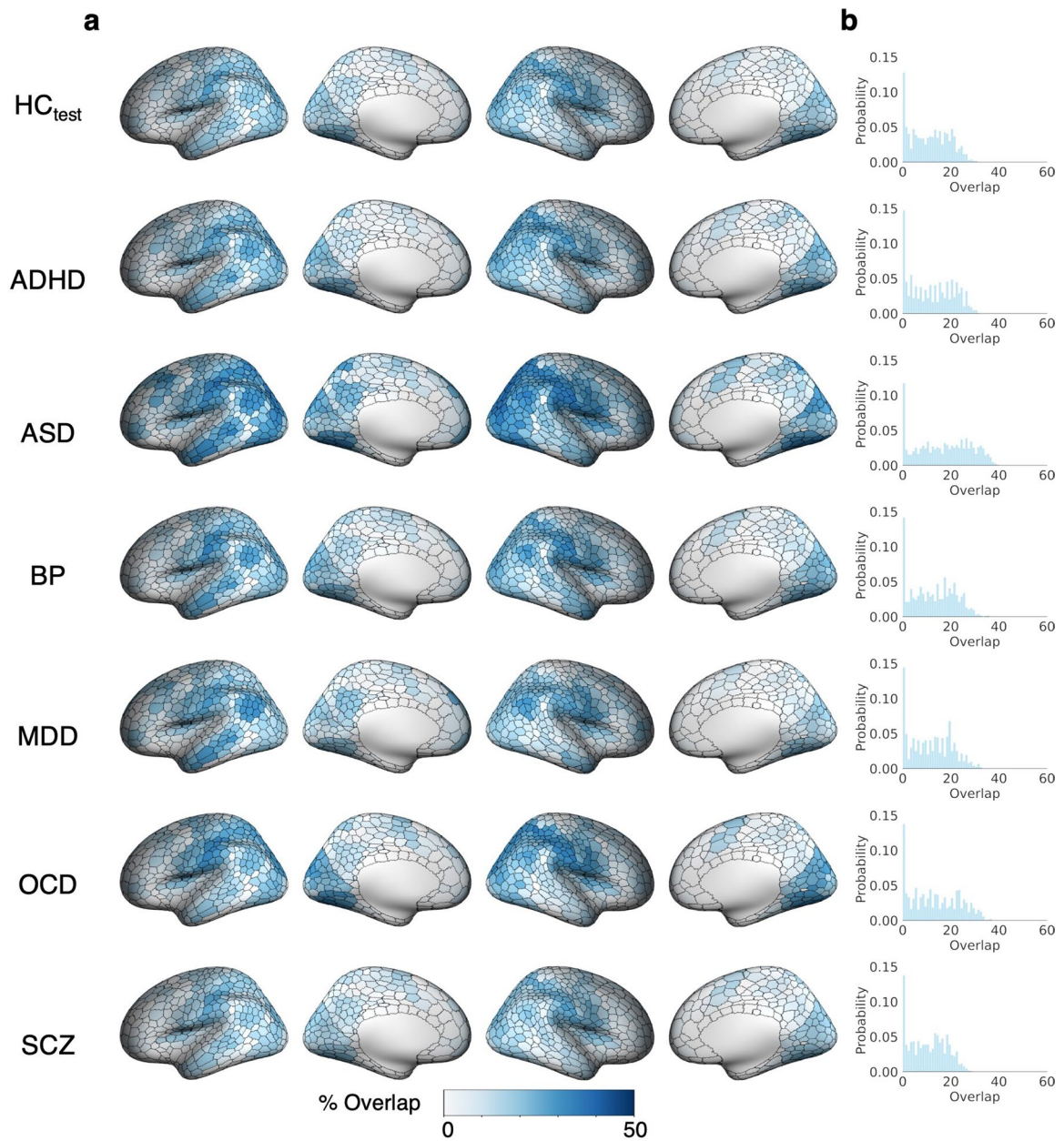


Extended Data Fig. 6 | Spatial overlap of extreme positive GMV deviations ($Z > 2.6$) in each group. a) Cortical and subcortical surface renderings showing spatial of overlap in 1032 brain regions, and b) the distribution of overlap percentages observed across all regions. Data used to generate this figure can be found in Supplementary Data 1 (*Regional_pos_thr26*).



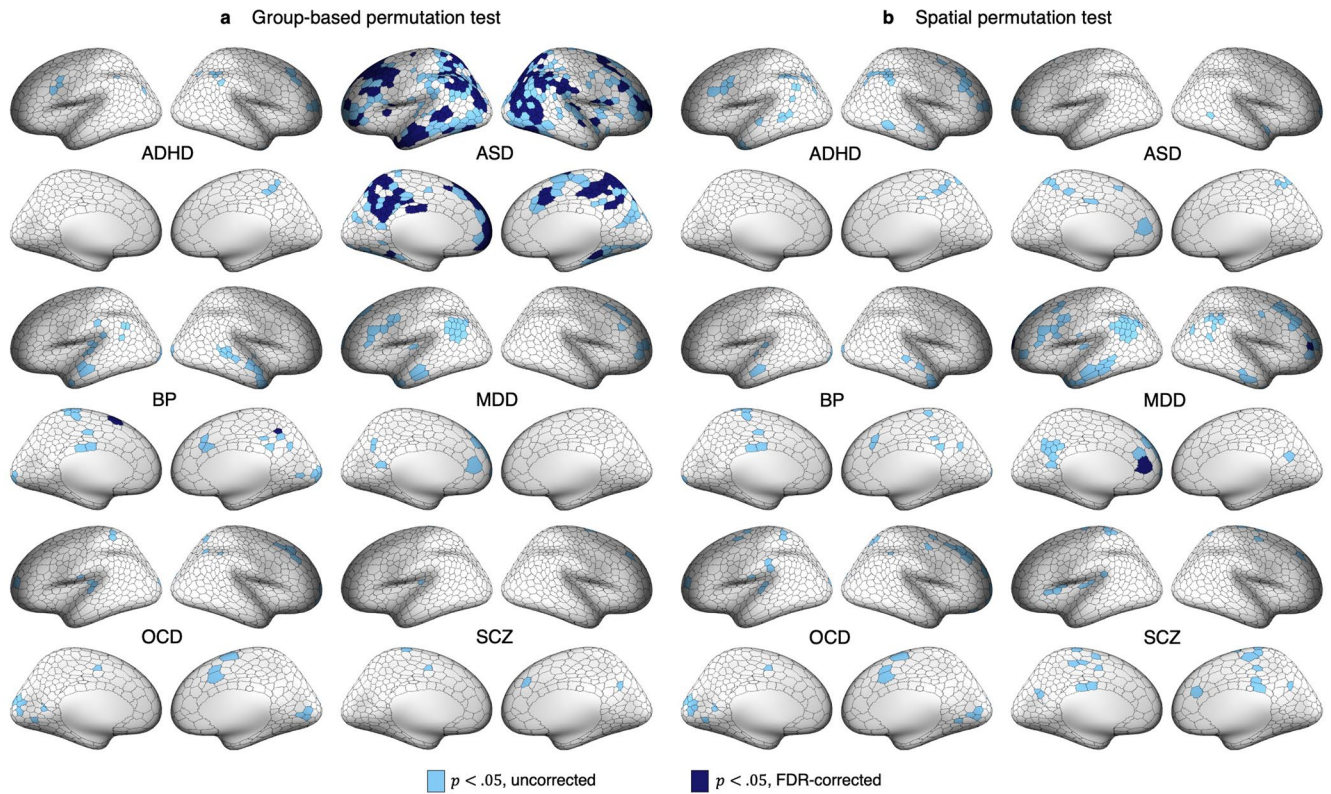
Extended Data Fig. 7 | Regional heterogeneity of extreme positive GMV deviations in each disorder. Cortical and subcortical surface renderings showing regions with significantly greater overlap of extreme positive GMV deviations in cases compared to controls identified using group-based

permutation tests (light blue corresponds to $p_{\text{uncorrected}} < .05$, dark blue corresponds to $p_{\text{FDR}} < .05$, two-tailed, cases>controls). Data used to generate this figure can be found in Supplementary Data 1 (*Regional_pos_thr26*).



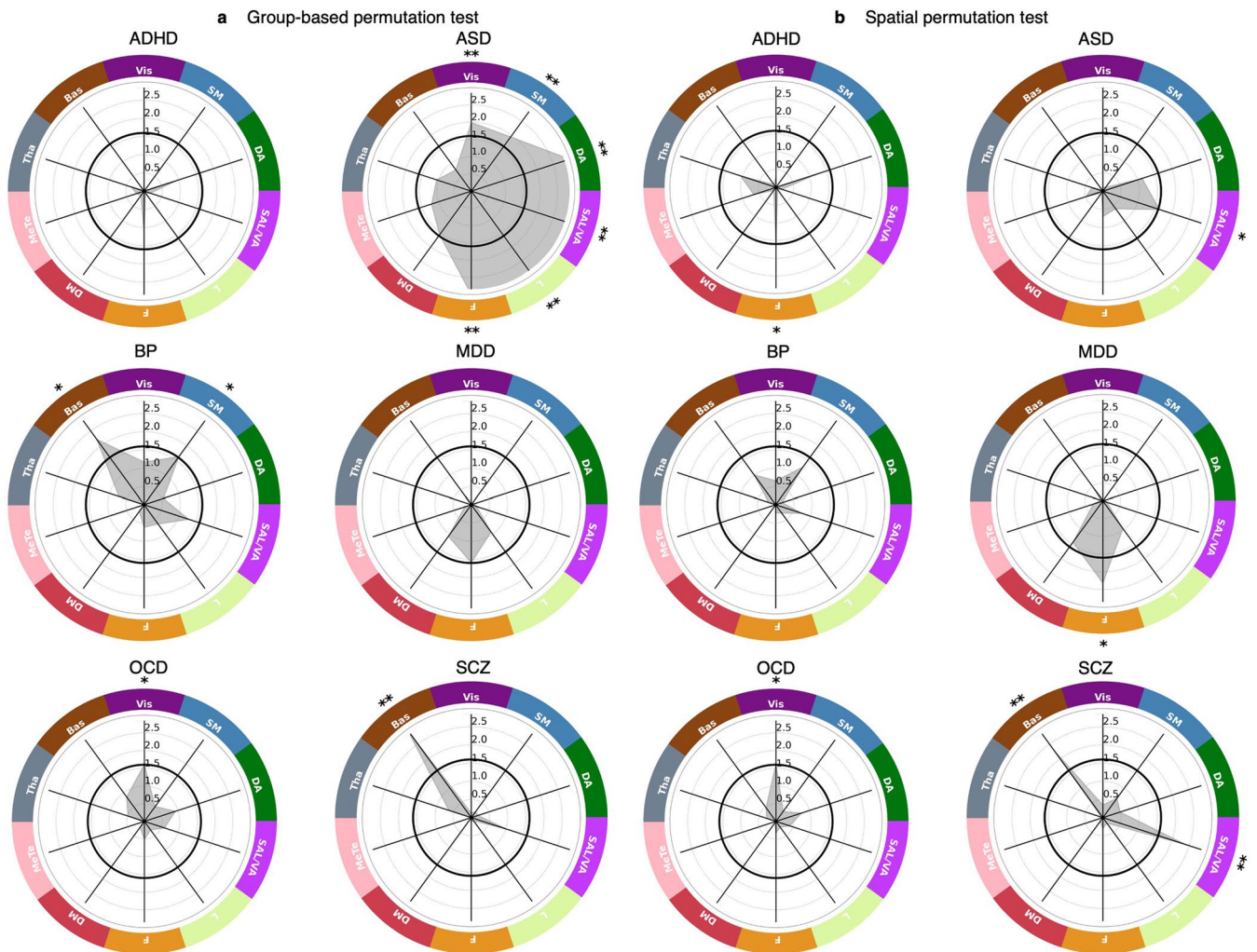
Extended Data Fig. 8 | Spatial overlap in regions functionally coupled (vertex-wise threshold $p_{FWE} < 0.025$), to extreme positive deviations ($Z > 2.6$) across groups, using a parcel-mapping threshold of 50%). a) Cortical

surface renderings showing spatial overlap, and b) the distribution of overlap percentages observed across all regions. Data used to generate this figure can be found in Supplementary Data 1 (*Circuit_pos_parc50*).



Extended Data Fig. 9 | Regions showing greater overlap in areas functionally coupled to extreme positive GMV deviations in cases compared to controls. Group differences in circuit-level overlap were evaluated with respect to two empirical null models (see Extended Data Fig 3 for details). (a) and (b) show cortical surface renderings of regions with significantly greater overlap in cases

compared to controls in areas functionally coupled to extreme deviations identified using group-based or spatial permutation tests, respectively (light blue corresponds to $p_{uncorrected} < .05$, dark blue corresponds to $p_{FDR} < .05$, two-tailed, cases < controls). Data used to generate this figure can be found in Supplementary Data 1 (*Circuit_pos_parc50*).



Extended Data Fig. 10 | Functional networks showing greater overlap in extreme positive GMV deviations in cases compared to controls. The network-level $-\log_{10}$ p-values associated with difference in percent overlap for extreme positive GMV deviations between each clinical group and the HC_{test} cohort. ** corresponds to $p_{FDR} < 0.05$, two-tailed, cases > controls, * corresponds

to $p_{uncorrected} < 0.05$, two-tailed, cases > controls. The solid black line indicates $-\log_{10} p = 1.6$ ($p = 0.05$, two-tailed, uncorrected). (a) and (b) identify networks showing significant differences under group-based or spatial permutation testing, respectively. Data used to generate this figure can be found in Supplementary Data 1 (*Network_pos_10network*).

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted <i>Give P values as exact values whenever suitable.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

| | |
|-----------------|--|
| Data collection | No software was used for data collection. Neuroimaging and behavioural data were from existing datasets (detailed below) whose acquisition's are presented in detail in previous work. |
| Data analysis | <p>T1-weighted MRIs were preprocessed using the default settings of the Computational Anatomy Toolbox (CAT12, r113, http://dbm.neuro.uni-jena.de/cat/) and Statistical Parametric Mapping (SPM12, http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), using MATLAB v9.8. This involved using Diffeomorphic Anatomical Registration Exponentiated Lie Algebra (DARTEL) to normalise the segmented scans. Example code and preprocessing report can be found here: https://github.com/ashlea-segal/multiscale-heterogeneity-brain-abnormalities</p> <p>rs-fMRI HCP data was preprocessed using the HCP minimal preprocessing pipeline (Glasser et al., 2013, Neuroimage). This included FSL FLIRT, TOPUP, ICA-FIX, and HCP Connectome Workbench (v1.2.3). HCPpipelines TaskfMRIAnalysis was adapted for the rs-fMRI seed-based analysis. Code can be found here: https://github.com/Washington-University/HCPpipelines.</p> <p>Normative Modelling Software: PCNtoolkit (version= 0.16, https://github.com/amarquand/PCNtoolkit) to generate person-specific z-score deviation maps.</p> <p>Permutation Analysis of Linear Models software package (PALM alpha116, https://github.com/andersonwinkler/PALM) for statistical inference</p> <p>Specific packages used within Anaconda virtual environment and code to generate figures can be found here: https://github.com/ashlea-segal/multiscale-heterogeneity-brain-abnormalities</p> |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

This study used neuroimaging data collected across multiple independent datasets, with varying levels of data accessibility. Below summarizes the availability of each dataset.

Autism Brain Imaging Data Exchange I (ABIDE I) and ABIDE II data used in this study came from the ABIDE 1000 Functional Connectomes Project repository, http://fcon_1000.projects.nitrc.org/indi/abide/. Data usage is unrestricted for non-commercial research purposes. As per INDI protocol, users must register with the NITRC and 1000 Functional Connectomes Project to gain access.

The Australian Schizophrenia Research Bank (ASRB) is a medical research database and storage facility that links clinical and neuropsychological information, blood samples and structural MRI brain scans from people with schizophrenia and healthy non-psychiatric controls. Data is available in the ASRB repository, subject to approval of the ASRB Access Committee <https://www.neura.edu.au/discovery-portal/asrb/>

First Episode Mania Study (FEMS) - Available from the principal investigator of the study, subject to local ethics committee requirements.

Monash Cohort (MON) - Available from the principal investigator of the study, subject to local ethics committee requirements

International Multi-centre persistent ADHD CollaboraTion (IMpACT-NL) - Not publicly available due to privacy or ethical restrictions.

OpenNeuro - Kansas Musical Depression Study (KANMDD) - Available in the OpenNeuro repository, <https://openneuro.org/datasets/ds000171>

OpenNeuro - Massachusetts Institute of Technology Autism Study (MITASD) - Available in the OpenNeuro repository, <https://doi.org/10.18112/openneuro.ds000212.v1.0.0>

Obsessive-compulsive and problematic gambling study (OCDPG) - Available from the principal investigator of the study, subject to local ethics committee requirements

OpenNeuro - Russia fMRI Depression Study (RUSMDD) - Available in the OpenNeuro repository, <https://doi.org/10.18112/openneuro.ds002748.v1.0.5>

SPAINOCD - Available from the principal investigator of the dataset, subject to local ethics committee requirements

TOP15 - The data are not publicly available due to privacy or ethical restrictions.

OpenNeuro – University of Washington ASD Study (WASHASD) - Available in the OpenNeuro repository, <https://doi.org/10.18112/openneuro.ds002522.v1.0.0>

YoDA - Available from the principal investigator of the study, subject to local ethics committee requirements

Human Connectome Project (HCP) - Available in the Human Connectome Project repository, <https://www.humanconnectome.org/study/hcp-young-adult>. Users must agree to data use terms for the HCP before being allowed access to the data and ConnectomeDB, details are provided at <https://www.humanconnectome.org/study/hcp-young-adult/data-use-terms>.

No new data were collected for this manuscript. Across all datasets, we did not use any specific software for downloading the data. For details on data collection for each data, please see relevant references in Supplementary Table 1.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

No sample size calculations were performed. This projects uses open-access data and data made available through collaborators. Combined, the sample consists of 1465 healthy controls (HC; 54.47% male) and 1294 cases, taken from 14 different studies and 25 different scan sites. The clinical sample comprised 202 individuals with ASD (100% male), 153 individuals with ADHD (41.18% male), 228 individuals with BP (47.37% male), 161 individuals with MDD (34.16% male), 167 individuals with OCD (50.30% male), and 383 participants with SCZ (62.14% male). See manuscript for extended discussion.

| | |
|-----------------|---|
| | HCP data was used because of the large number of subjects and the relatively equal male/female balance. For HCP data, only unrelated subjects were included. |
| Data exclusions | <p>All T1-weighted images were visually inspected and evaluated for the presence of artifacts, resulting in the exclusion of 53 images with gross artifacts or abnormalities. Next, we used the Computational Anatomy Toolbox (CAT12 r113, http://dbm.neuro.uni-jena.de/cat/) to generate a weighted overall image quality rating (IQR) for every scan. This metric combines ratings of basic image properties, including the level of noise and geometric distortions, into a single score that quantifies the overall image quality of a participant's T1-weighted scan (see here for more information www.neuro.uni-jena.de/cat/index.html#QA). On this metric, lower scores denote higher image quality. As per previous work²¹, we excluded 153 images with an IQR > 2.8. An additional 63 images were excluded due to a failure of our processing pipeline. Collectively, this quality assurance process resulted in the exclusion of a total of 269 scans.</p> <p>Following this quality control protocol, we performed a series of additional quality control checks and exclusions for our analysis. Specifically, we excluded participants if they were below 18 years or above 64 years of age (N=346); did not have the necessary clinical data (clinical diagnosis or, for healthy controls, absence of any clinical diagnosis) or demographic information (age, sex and scanner site, N=53); if their T1-weighted structural magnetic resonance imaging (MRI) scan did not survive our stringent manual and automated quality control procedure, as explained below (N=269); or if the data came from a site with less than 10 individuals in the same group and sex (described in the Normative model section, below; N=217). Our final sample available for analysis thus comprised 1465 HCs and 1294 cases.</p> |
| Replication | <p>We have included numerous quality control checks to verify the stability of our findings, including:</p> <ul style="list-style-type: none"> • Cross-validation of the normative models to assess model generalisability • Multiple linear support vector machines to evaluate the model's efficacy in partitioning site-related variance • Pearson's correlation between number of extreme deviations and image quality to assess whether individual variations in scan quality affected the deviation zmaps • Threshold-weighted analysis of deviation mapping to show that our results do not depend on the choice of a specific threshold for defining deviations • The inclusion of a held-out control group for all analyses to establish a normative benchmark for assessing % overlap in the clinical groups • Mapping circuit-level overlap using two mapping thresholds (50% and 75%) to ensure our findings were not driven by this specific choice • Statistical inference at the circuit- and network- level using two different null models to disentangle whether any differences arise from a preferential aggregation of deviations or whether group differences are driven by variations in total deviation burden |
| Randomization | Randomization was not performed because participants were not placed into experimental groups - case and control groups are defined based on diagnosis. Participants without a diagnosis of a psychiatric disorder were in the control group. Participants diagnosed with a psychiatric disorder were placed in the clinical group, which was further divided by diagnosis (ADHD, ASD, BP, OCD, MDD, SCZ). |
| Blinding | Blinding is not relevant to this study because participants were not placed into experimental groups. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

| n/a | Involved in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> MRI-based neuroimaging |

Human research participants

Policy information about [studies involving human research participants](#)

| | |
|----------------------------|---|
| Population characteristics | See above in Life sciences study design: Sample size. All participants were aged between 18-64 years old. |
| Recruitment | See the relevant reference for each dataset for specific details of each dataset. In general, the sample was collected through a range of different recruitment resources including media advertisements, inpatient, outpatient and community mental health service providers, non-government organizations, rehabilitation services, and cold telephoning using the electoral rolls. |
| Ethics oversight | The study was approved by the local ethics committee of each dataset, and written informed consent was obtained from each participant. This study was approved by the Monash University Research Ethics Committee (Project ID: 23534). |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

| | |
|---------------------------------|---|
| Design type | resting-state fMRI, structural (VBM) MRI |
| Design specifications | HCP: The rs-fMRI data for each subject consists of four runs (rfMRI_REST1_LR, rfMRI_REST1_RL, rfMRI_REST2_LR, rfMRI_REST2_RL) acquired at two different sessions (REST1 and REST2) using two different directions of phase coding (LR: left to right, and RL: right to left). |
| Behavioral performance measures | Behavioral performance measures were not applicable in this study |

Acquisition

| | |
|-------------------------------|---|
| Imaging type(s) | resting-state fMRI, structural (VBM) MRI |
| Field strength | Described in Supplementary Table 1 for each dataset. |
| Sequence & imaging parameters | Described in Supplementary Table 1 for each dataset for structural data. HCP rs-fMRI data were obtained using a 32-channel Siemens 3T connectome-Skyra scanner. The imaging parameters for rs-fMRI were as follows: repetition time (TR)=720 ms, echo time (TE)=33.1 ms, flip angle (FA)=52°, field of view (FOV)=208×180 mm ² , matrix=104×90, slice number=72, slice thickness=2 mm, voxel size=2×2×2 mm ³ , multiband factor=8, and 1200 volumes. |
| Area of acquisition | Whole brain |
| Diffusion MRI | <input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used |

Preprocessing

| | |
|----------------------------|--|
| Preprocessing software | Preprocessing of the structural and functional data was done using a suite of tools. Structural preprocessing GMV was estimated using the CAT12 VBM pipeline, which is included as an extension of Statistical Parametric Mapping software (SPM12, http://www.fil.ion.ucl.ac.uk/spm/software/spm12) in MATLAB v9.8 using the default settings. Briefly, the T1-weighted images were first corrected for intensity nonuniformities, and segmented into GM, white matter (WM), and cerebrospinal fluid (CSF) tissue probability maps. Using the high-dimensional Diffeomorphic Anatomical Registration Exponentiated Lie Algebra (DARTEL), the segmented scans were normalized into standard Montreal Neurological Institute (MNI) space. Lastly, the images were bias-field corrected and modulated by the linear and nonlinear components of the Jacobian determinant obtained from the DARTEL deformation fields to obtain voxel-wise estimates of GMV. Functional preprocessing The rs-fMRI data used underwent the HCP's minimal pre-processing pipeline, which includes gradient-nonlinearity-induced distortion, motion correction to the single-band reference image using FLIRT, EPI image distortion correction using TOPUP, registration into standard space using a customized boundary-based-registration (BBR) algorithm, and single step spline interpolation using all transforms, intensity normalization and bias field removal to resample the original EPI into MNI space. Minimal high pass filtering was applied with a cutoff of 2000ms. Artifacts were then removed using ICA-FIX. This involves employing an automatic classifier, specifically trained for HCP data, to identify ICA components due to measurement noise, additional motion or physiological artifacts like cardiac pulsation and respiration. Next, the volume timeseries were mapped into the standard CIFTI grayordinate space and smoothed to 2mm FWHM (where the smoothing was on the surface for the cortex and in volume space for subcortex). This results in a standard set of grayordinates in every subject, with surface vertex data and subcortical volume voxel data. The mean GM signal was then removed each grayordinate's time series to remove residual widespread signal deflections that are not removed by ICA-FIX |
| Normalization | See above |
| Normalization template | MNI Standard Space |
| Noise and artifact removal | See above |
| Volume censoring | See above |

Statistical modeling & inference

| | |
|-------------------------|--|
| Model type and settings | Mass univariate across multiple brain scales (regional, circuit, network; see manuscript) |
| Effect(s) tested | As the primary aim of this study was to investigate anatomically heterogeneous brain deviations across multiple brain scales (regional, circuit, network) across six psychiatric disorders, multiple effects were examined. The main effect - observed group |

differences between each patient group and controls was evaluated with respect to two complementary nulls detailed in Supplementary Figure 6.

Specify type of analysis: Whole brain ROI-based Both

Anatomical location(s) Parcel-level analyses utilised the Schaefer et al., 2018 Cerebral Cortex 1000 region parcellation for cortical regions, and the Tian et al Nature Neuroscience 32 region parcellation for subcortical regions.

Statistic type for inference (See [Eklund et al. 2016](#)) Non-parametric inference was used across multiple neuroanatomical scales including regions of interest, and networks (see manuscript)

Correction Results were presented at $p < 0.05$ uncorrected and following False Discovery Rate (FDR, $q = 0.05$) correction.

Models & analysis

n/a | Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity Pearson correlation