Title

Networks extracted from nonlinear fMRI connectivity exhibit unique spatial variation and enhanced sensitivity to differences between individuals with schizophrenia and controls.

Permalink

Journal

Authors

Kinsey, Spencer
Kazimierczak, Katarzyna
Camazón, Pablo
et al.

Publication Date

2024

DOI

Peer reviewed

Article

Networks extracted from nonlinear fMRI connectivity exhibit unique spatial variation and enhanced sensitivity to differences between individuals with schizophrenia and controls

Check for updates

Spencer Kinsey [1,2] ✉, Katarzyna Kazimierczak[3], Pablo Andrés Camazón [4], Jiayu Chen[1], Tülay Adali[5], Peter Kochunov[6], Bhim M. Adhikari[6], Judith Ford[7,8], Theo G. M. van Erp[9], Mukesh Dhamala[1,2,10], Vince D. Calhoun [1,2,11] & Armin Iraji[1,2,11] ✉

Schizophrenia is a chronic brain disorder associated with widespread alterations in functional brain connectivity. Although data-driven approaches such as independent component analysis are often used to study how schizophrenia impacts linearly connected networks, alterations within the underlying nonlinear functional connectivity structure remain largely unknown. Here we report the analysis of networks from explicitly nonlinear functional magnetic resonance imaging connectivity in a case–control dataset. We found systematic spatial variation, with higher nonlinear weight within core regions, suggesting that linear analyses underestimate functional connectivity within network centers. We also found that a unique nonlinear network incorporating default-mode, cingulo-opercular and central executive regions exhibits hypoconnectivity in schizophrenia, indicating that typically hidden connectivity patterns may reflect inefficient network integration in psychosis. Moreover, nonlinear networks including those previously implicated in auditory, linguistic and self-referential cognition exhibit heightened statistical sensitivity to schizophrenia diagnosis, collectively underscoring the potential of our methodology to resolve complex brain phenomena and transform clinical connectivity analysis.

Schizophrenia is a brain disorder thought to be underpinned by altered neural interactions at various spatial and temporal scales[1]. At the whole-brain level, functional magnetic resonance imaging (fMRI) functional connectivity (FC) analysis is a non-invasive approach that has commonly been used to study how schizophrenia-related brain alterations are reflected within statistical relationships between blood-oxygenation-level-dependent (BOLD) time series. Although the relationship between the BOLD signal and neural activity is indirect[2], experimentally induced and resting-state BOLD fluctuations are typically associated with changes in local field potentials across multiple frequency bands[3–6], indicating that fMRI FC analysis is a promising tool for advancing the identification of task-related and spontaneously emerging networks of interacting brain regions. Moreover, fMRI FC analysis is deployable within a wide range of predictive clinical contexts. For example, multiple large-scale meta-analyses have shown that FC measures reliably distinguish healthy controls (HC) from individuals

A full list of affiliations appears at the end of the paper. ✉e-mail: skinsey8@gsu.edu; airaji@gsu.edu

with schizophrenia (SZ)[7–9]. Such studies have contributed to an accumulation of evidence for the SZ 'dysconnection hypothesis'[10], which posits FC alteration as a central endophenotype of the disorder resulting from neuromodulatory and synaptic pathogenesis.

However, FC studies are typically designed to estimate networks that reflect linear statistical relationships between brain areas[11–13]. Although the remarkably complex nonlinear interactions inherent to brain networks have been recognized and investigated[14–18], there is a need to develop data-driven methods capable of estimating networks that accurately reflect the structure of these nonlinear connectivity patterns[19] and thus fill the gap in knowledge concerning their contributions to brain function and alterations in psychiatric disorders such as SZ. In this Article we highlight three ways in which decomposing nonlinear brain connectivity patterns into data-driven networks has the potential to advance systems, cognitive and predictive clinical neuroscientific research. First, effectively capturing networks from nonlinear patterns in a data-driven fashion may lead to a more precise and thorough characterization of the organization and dynamics of neural ensembles at multiple scales[16,20,21]. Second, networks that accurately reflect underlying nonlinear connectivity patterns may reveal unique associations with cognitive and behavioral capacities. In principle, nonlinearity is thought to underpin a high-dimensional state space capable of supporting a set of flexible and diverse neural computations[15,17], such that analyzing the functional role of nonlinear encoding of information[22] may shed light on the structure of cognitive processes and deficiencies associated with psychiatric disorders such as SZ and their symptoms. Third, networks captured from measures that are sensitive to nonlinearity can be leveraged to develop biomarkers that can be incorporated within brain-based predictive models of mental illness, or 'predictomes'[23].

Among the available network estimation methods, independent component analysis (ICA) is known to be a powerful multivariate source separation technique[24,25]. ICA assumes that the data are a linear mixture of statistically independent source signals and aims to estimate an unmixing matrix, yielding components that approximate these signals optimally[24,25]. In the context of fMRI FC analysis, spatial ICA has commonly been used to decompose fMRI time-series data into a set of intrinsic connectivity networks (ICNs), where the spatial pattern of a network describes its distribution across voxels and the temporal pattern describes its activity over time[11,26–28]. ICNs can be robustly and consistently identified from both resting-state[26,27,29] fMRI (rsfMRI) and task-based[11,26,30,31] fMRI (tfMRI) time-series data at different spatial scales[20,21,32]. ICNs can also be reliably extracted from FC matrices constructed from second-order statistics such as Pearson correlation (that is, from the connectivity domain)[33,34]. Connectivity-domain ICA is a type of feature-based analysis[35] that yields cross-validating components, and it is distinguished from time-domain ICA by unique benefits such as consistency across changes in particular analysis parameters and reproducibility[33]. Moreover, an expanding range of FC metrics can be used to construct the connectivity basis, making connectivity-domain ICA an incredibly versatile tool[33].

Connectivity- and time-domain ICA have become valuable tools for investigating fMRI data. However, both methods are typically designed to identify ICNs composed of covarying brain regions, thereby capturing ensembles explained by linear connectivity information[11,33,34]. Although recent advancements have made strides in incorporating nonlinearity, such as learning local spatial or temporal nonlinear structures[36,37], the extent to which the estimated sources reflect nonlinear connectivity patterns remains unclear. To address this gap in knowledge, we advance an approach to extract ICNs from distance correlation[38] patterns that move beyond those constructed from Pearson correlation (Fig. 1). We first estimate explicitly nonlinear whole-brain FC (ENL-wFC) by using a linear regression-based approach to remove the nonlinear whole-brain FC information (NL-wFC;
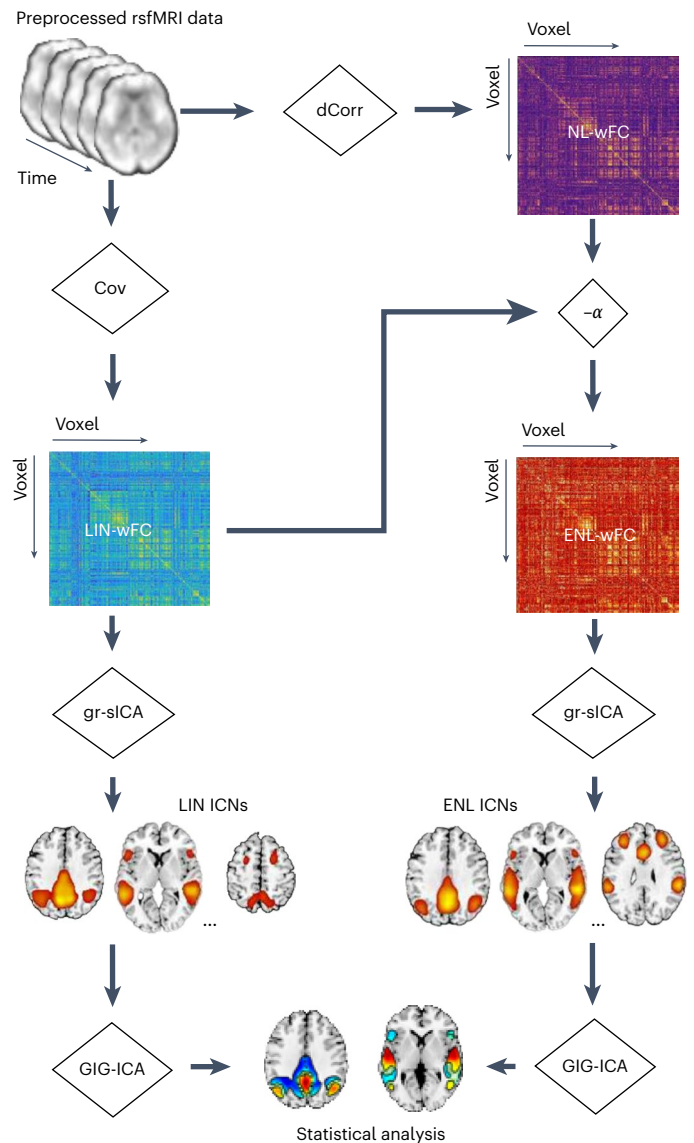

Fig. 1 | Schematic of the analysis pipeline. Preprocessed rsfMRI data are transformed to the connectivity domain using Cov (as a linear FC estimator) and dCorr as a nonlinear FC estimator. ENL-wFC is obtained by removing the NL-wFC information which is linearly explained by LIN-wFC. Gr-sICA is implemented in the connectivity domain on LIN-wFC and ENL-wFC to estimate separate sets of intrinsic connectivity networks (LIN and ENL ICNs). GIG-ICA is then used to estimate subject-specific ICNs, and statistical analysis is conducted on the subject-level spatial maps.

operationalized as distance correlation) explained by linear whole-brain FC (LIN-wFC), and we subsequently implement group-level spatial ICA (gr-sICA) in the connectivity domain[39], resulting in a targeted analysis of network features that are inaccessible to approaches that aim to compute brain connectivity using methods that incorporate both linear and nonlinear information. Although alternate metrics can be used to quantify fMRI connectivity while accounting for higher-order statistics[40–42], distance correlation is a powerful and flexible dependence metric that remains underexplored in the context of FC research. Moreover, the proposed method is unique, in that we conceive of ENL-wFC as a global feature of the connectivity space rather than as a composite feature constructed from pairwise associations[41,42]. This allows us to leverage information present within global connectivity features beyond those found within macroscopic linear connectivity patterns. In this Article we use this approach to assess
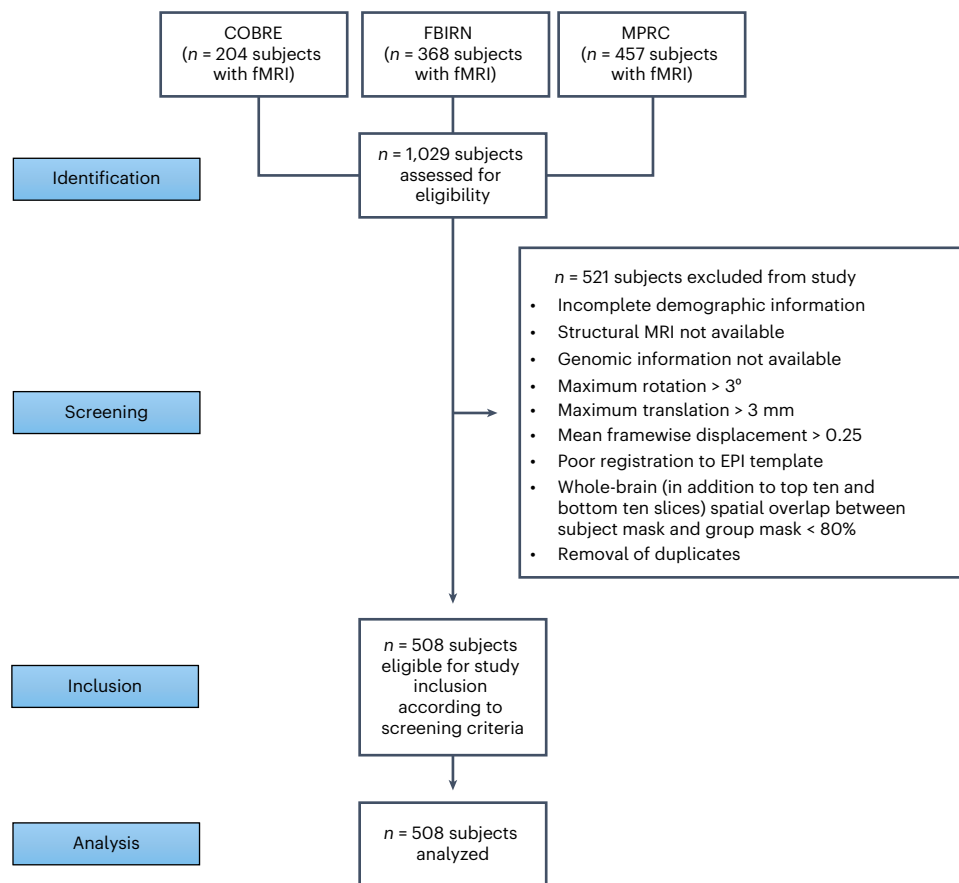
Fig. 2 | Strengthening the reporting of observational studies in epidemiology (STROBE) flowchart. COBRE, Center for Biomedical Research Excellence; FBIRN, Functional Imaging Biomedical Informatics Research Network; MPRC, Maryland Psychiatric Research Center. EPI, echo-planar imaging.

Table 1 | Subject demographic information

Dataset	Diagnosis (no.)	Sex (no.)	Race (AMR/EUR/AFR/other)	Age (years)[a]	Age (years)[b]
COBRE	HC (75)	Male (56)	20/32/4/0	39.27±12.13	39/(18–65)
		Female (19)	13/3/3/0	35.47±10.02	34/(18–58)
	SZ (51)	Male (45)	19/23/3/0	37.36±15.28	33/(19–64)
		Female (6)	4/1/1/0	40.83±17.70	44/(20–65)
FBIRN	HC (88)	Male (60)	12/48/0/0	36.58±10.74	39/(19–59)
		Female (28)	6/22/0/0	36.61±11.07	33/(19–58)
	SZ (60)	Male (52)	17/35/0/0	39.54±11.12	41/(18–60)
		Female (8)	1/7/0/0	35.88±9.85	34/(24–51)
MPRC	HC (152)	Male (69)	6/43/19/1	38.99±13.22	41/(18–68)
		Female (83)	8/44/30/1	39.93±14.93	43/(16–64)
	SZ (82)	Male (57)	3/33/19/2	36.25±13.54	33/(13–63)
		Female (25)	0/13/11/1	44.68±11.92	47/(13–61)

COBRE, Center for Biomedical Research Excellence. FBIRN, Functional Imaging Biomedical Informatics Research Network. MPRC, Maryland Psychiatric Research Center. HC, healthy control. SZ, schizophrenia. AMR, mixed American. EUR, European. AFR, African. [a]Mean±s.d. [b]Median/range.

differences in spatial variation between explicitly nonlinear (ENL) and linear (LIN) network estimates and to investigate SZ-associated network alterations in a multi-study rsfMRI dataset sourced from three major psychosis projects: the Center for Biomedical Research Excellence (COBRE)[43], the Functional Imaging Biomedical Informatics Research Network (FBIRN)[44,45] and the Maryland Psychiatric Research Center (MPRC) (Fig. 2 and Table 1)[46].

Results
Goodness of fit
Goodness-of-fit statistics (R^2) for the linear regression of NL-wFC on LIN-wFC were as follows: mean ± s.d. = 0.5337 ± 0.2009; minimum–maximum = 0.0173–0.9413. This indicates that, on average, much of the NL-wFC variance is captured by a linear fit. After accounting for confounding factors (Methods), R^2 is significantly higher for

HC versus SZ ($n = 508$; $P = 0.0002$, observed difference = 0.1203, Hedges's $g = 0.6971$). The observed Hedges's g value indicates the presence of a medium to large effect size. HC residual indices were mean ± s.d. = 0.0457 ± 0.1660; minimum−maximum = −0.4912−0.4936. SZ residual indices were mean ± s.d. = −0.0746 ± 0.1827; minimum−maximum = −0.4725−0.3800.

Component estimation reliability is greater for ENL versus LIN

Components estimated from ENL-wFC exhibit significantly higher estimation reliability (ICASSO IQ) compared to components estimated from LIN-wFC ($n = 40$ components; $P = 0.0006$, observed difference = 0.037, Hedges's $g = 0.6441$). The observed Hedges's g-value indicates the presence of a medium to large effect size. ENL stability indices were mean ± s.d. = 0.9694 ± 0.0057; minimum−maximum = 0.9579−0.9800. LIN stability indices were mean ± s.d. = 0.9324 ± 0.0810; minimum−maximum = 0.6186−0.9770.

Common and unique ICNs identified from ENL and LIN

Within our 20-model-order gr-sICA framework, 13 ENL ICNs and 14 LIN ICNs were identified (Fig. 3). Among the identified networks, ten exhibited maximum spatial similarity values exceeding 0.80 between their ENL and LIN estimates. We classified these networks as common to both ENL-wFC and LIN-wFC based on the defined criterion (Methods). Among those remaining, two ENL and three LIN ICNs exhibited maximum spatial similarity values between 0.40 and 0.80. Although several of these networks attained relatively high maximum spatial similarity, we noticed distinct intensity differences across their neuroanatomical distributions that prevented common classification and labeling. Furthermore, our analysis uncovered a LIN network and an ENL network exhibiting a maximum spatial similarity less than 0.40. We classified these networks as unique based on our uniqueness criterion (Methods), and we validated the uniqueness of the ENL ICN in question across 100 additional iterations of gr-sICA (Supplementary Note 1 and Supplementary Fig. 1).

ENL and LIN ICNs exhibit unique spatial patterns

ENL and LIN ICNs exhibit distinctive spatial distributions (Fig. 4a–j). Visible gradients are present within networks associated with both lower and higher cognitive functioning, and many core regions (defined as regions that attain higher values across the spatial distribution) exhibit greater ENL weight. For the subcortical (SUB) network (Fig. 4a), LIN weight is greater within the bilateral caudate and putamen, and ENL weight is greater within the bilateral thalamus. The cerebellum (CER; Fig. 4b) exhibits higher ENL weight within vermis lobules I–V, and higher LIN weight within lobules VII–IX and the bilateral hemisphere. Among networks associated with visual[47] and auditory and linguistic[48] functioning, ENL weight is predominantly greater within spatially central regions, whereas LIN weight is greater within peripheral areas. For instance, the primary visual (VIS1) network (Fig. 4c) exhibits a medial–lateral gradient in the bilateral cortex surrounding the calcarine fissure, with greater ENL weight within the cuneus. The secondary visual (VIS2) network (Fig. 4d) shows higher ENL weight within the cuneus and higher LIN weight within the bilateral inferior and middle occipital gyri. Temporal (TEMP) network (Fig. 4e) variation follows a similar center–periphery pattern, with greater ENL weight in the superior temporal gyri and greater LIN weight within the supramarginal gyri and bilateral inferior frontal triangularis.

Whereas both the primary and secondary sensorimotor networks (MTR1 and MTR2) exhibit gradients (Fig. 4f–g), MTR1 comparisons reveal a medial–lateral pattern between the paracentral lobules and pre- and postcentral gyri, while MTR2 comparisons reveal an inferior–superior gradient between the superior temporal lobe and pre- and postcentral gyri. Networks implicated in higher cognitive functions such as attention[49], social cognition and self-referential processes[50], and executive control[51] exhibit core–periphery gradients. The dorsal
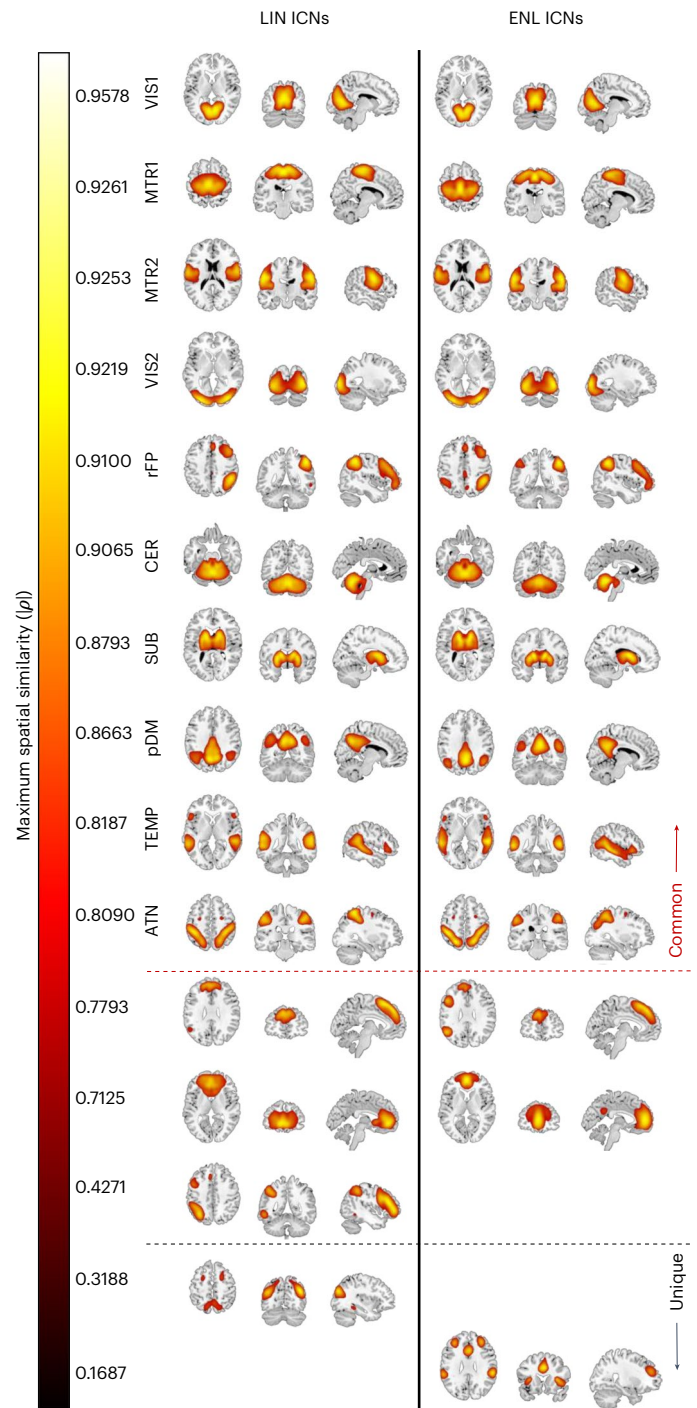

Fig. 3 | ICNs obtained from LIN-wFC and ENL-wFC gr-sICA in the connectivity domain. ICNs are displayed using an empirical threshold ($Z > 1.96$; $P < 0.05$) on the ch2bet template in order of maximum spatial similarity. Common ICNs (maximum similarity > 0.80) include primary visual (VIS1), primary sensorimotor (MTR1), secondary sensorimotor (MTR2), secondary visual (VIS2), right frontoparietal (rFP), cerebellum (CER), subcortical (SUB), posterior default mode (pDM), temporal (TEMP) and dorsal attention (ATN). ICNs exhibiting maximum similarity between 0.40 and 0.80 and unique ICNs (maximum similarity < 0.40) are also displayed.

attention (ATN; Fig. 4h) network shows higher ENL weight in the superior parietal lobules and higher LIN weight in the postcentral gyri. The posterior default mode (pDM) network (Fig. 4i) exhibits higher ENL values in the precuneus and bilateral angular gyri, with
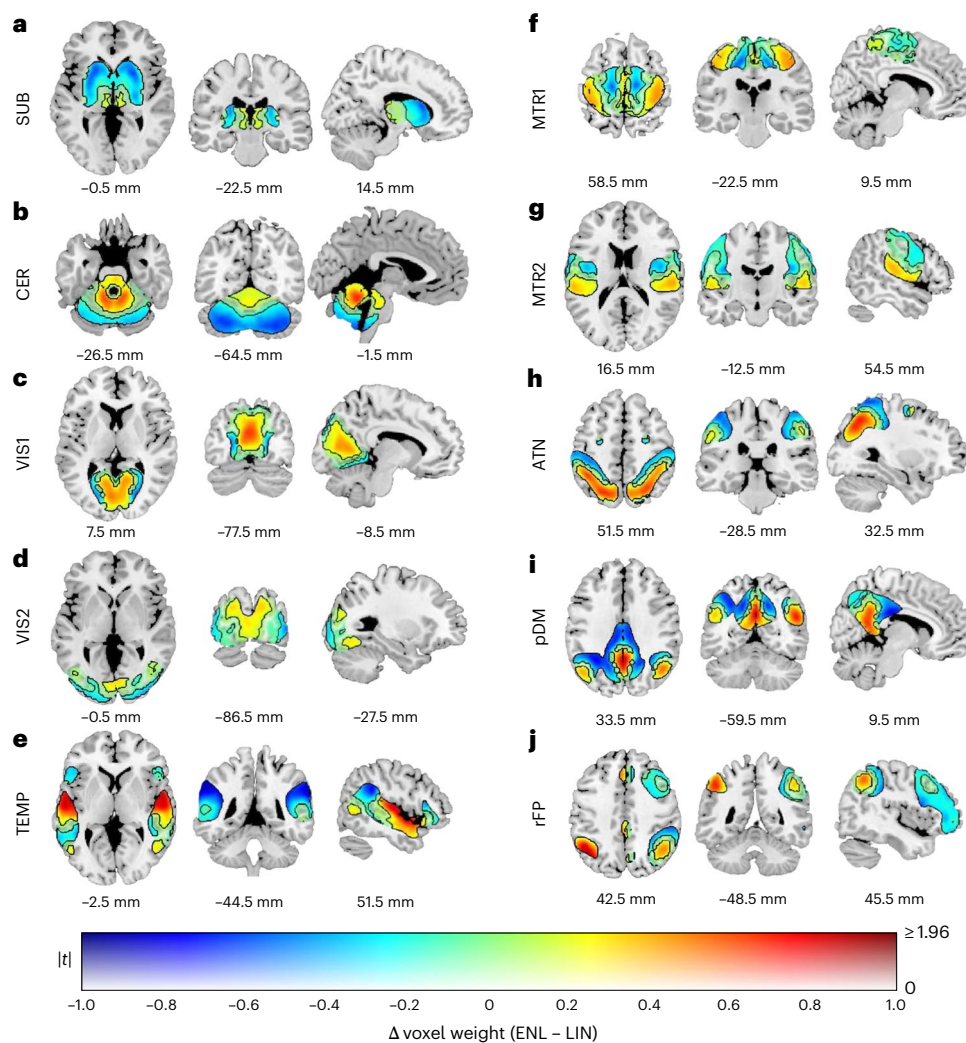
Fig. 4 | Assessment of ICN spatial variation. Results are plotted according to a dual-coded[86] colormap, with transparency reflecting two-sided paired sample *t*-statistic magnitudes and contours indicating FDR-corrected statistical significance ($q < 0.05$). Warmer hues indicate ENL > LIN, and cooler hues indicate LIN > ENL. **a–j**, The displayed ICNs are subcortical (SUB) (**a**), cerebellum (CER) (**b**), primary (VIS1) (**c**) and secondary (VIS2) (**d**) visual, temporal (TEMP) (**e**), primary (MTR1) (**f**) and secondary (MTR2) (**g**) sensorimotor, dorsal attention (ATN) (**h**), posterior default mode (pDM) (**i**) and right frontoparietal (rFP) (**j**). The results are overlaid on the ch2bet template with *x*, *y* and *z* coordinates listed relative to the origin in Montreal Neurological Institute 152 space.

higher LIN values in the middle and posterior cingulate. The right frontoparietal (rFP) network (Fig. 4j) exhibits higher ENL values within the angular gyri (particularly within the left angular gyrus) and higher LIN values within the right inferior parietal lobule, right middle frontal gyrus and right inferior frontal triangularis. The robustness of our voxel-wise *t*-test assessment of spatial variation was confirmed by permutation test results from pDM comparisons ($\rho = 0.9653$). Summary test information for the spatial variation analysis is provided in Supplementary Table 1.

ENL ICN voxels exhibit enhanced sensitivity to SZ diagnosis
Collectively, ENL network voxels exhibit a greater degree of sensitivity to SZ diagnosis versus LIN ($\chi^2 = 53.75$; $P < 0.00001$; odds ratio = 1.24), and a greater number of ENL voxels are implicated (Supplementary Table 2). Moreover, ENL counterparts of networks implicated in auditory and linguistic[48,52,53], sensorimotor[54] and self-referential[50] cognitive processes exhibit enhanced sensitivity to differences between HC and SZ (Fig. 5a–c). For example, although both sets of comparisons revealed differences within TEMP regions comprising the primary auditory and auditory association cortex, ENL comparisons are more sensitive ($\chi^2 = 851.3$; $P < 0.00001$; odds ratio = 22.63), revealing clusters that are

more numerous, with augmented volumes and effect sizes (Fig. 5a). LIN and ENL tests revealed higher values for HC within the bilateral superior temporal gyri and temporal poles, bilateral insula, bilateral Heschl's gyrus, bilateral Rolandic operculum and right middle temporal gyrus, along with higher values for SZ within the right supramarginal gyrus. However, ENL tests revealed a larger number of significant voxels across these regions. Additionally, ENL tests revealed higher HC values within the left middle temporal gyrus and higher SZ values within the left supramarginal gyrus, both of which were missed for significance by LIN tests.

ENL sensitivity was also greater for MTR2 (Fig. 5b; $\chi^2 = 639.5$; $P < 0.00001$; odds ratio = 7.61) and pDM (Fig. 5c; $\chi^2 = 125.03$; $P < 0.00001$; odds ratio = 128) tests. Both sets of MTR2 comparisons revealed greater weight for HC within the bilateral postcentral gyri. However, ENL tests revealed more extensive clusters and greater HC values within the bilateral posterior insula. ENL pDM tests revealed clusters of higher values for SZ within the precuneus and left angular gyrus, whereas LIN comparisons identified only two significant voxels. However, we note that LIN tests were more sensitive for CER ($\chi^2 = 445.7$; $P < 0.00001$; odds ratio = 3.93), VIS1 ($\chi^2 = 76.85$; $P < 0.00001$; odds ratio = 2.28), VIS2 ($\chi^2 = 391$; $P < 0.00001$; odds ratio = 394), MTR1
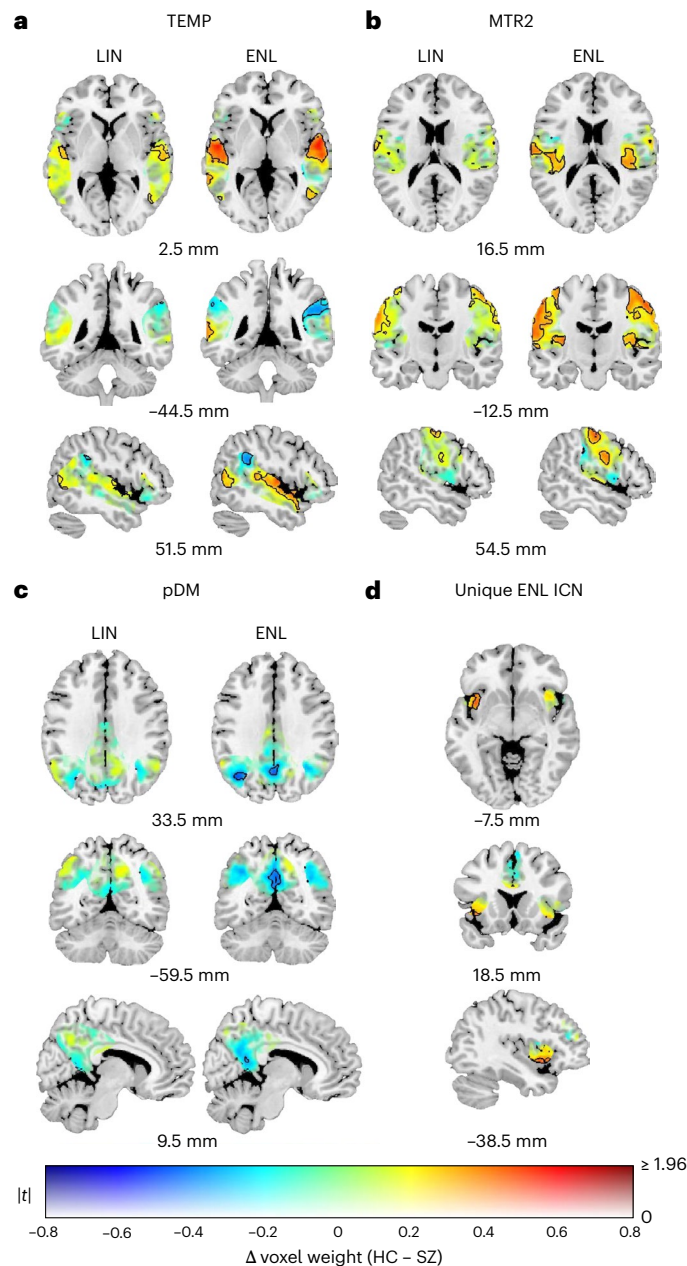
Fig. 5 | Statistical comparisons between subject-level ICN estimates from healthy controls and individuals with schizophrenia. a–d, Comparisons between subject-level temporal (TEMP) (**a**), secondary sensorimotor (MTR2) (**b**), posterior default mode (pDM) (**c**) and unique explicitly nonlinear (ENL) (**d**) ICN estimates derived from healthy controls (HC) and individuals with schizophrenia (SZ). Results are plotted according to a dual-coded colormap[86], with transparency reflecting two-sided independent-samples t-statistic magnitudes and contours indicating FDR-corrected statistical significance ($q < 0.05$). In **a–c**, results from LIN comparisons are located on the left, and results from ENL comparisons are located on the right. Warmer hues indicate HC > SZ, and cooler hues indicate SZ > HC. Results are overlaid on the ch2bet template with x, y and x coordinates listed relative to the origin in Montreal Neurological Institute 152 space.

($\chi^2 = 6.868$; $P = 0.0088$; odds ratio = 1.76) and rFP (exact binomial test; $P < 0.00001$; odds ratio undefined) networks. Unique LIN network comparisons failed to reveal any group differences (Supplementary Fig. 2h), but unique ENL network comparisons revealed a cluster within the left anterior insula that distinguished cohorts, with HC exhibiting greater values than SZ (Fig. 5d, Fig. 2 and Table 1). The robustness of our voxel-wise t-test assessment of cohort differences was confirmed

by permutation test results from LIN ($\rho > 0.999$) and ENL ($\rho > 0.999$) pDM comparisons. Results for SUB, CER, VIS1, VIS2, MTR1, ATN, rFP and unique LIN network cohort comparisons are depicted in Supplementary Fig. 2. Summary cohort test information is provided in Supplementary Table 2 and summary sensitivity test information in Supplementary Table 3.

To validate the detection of SZ alterations within TEMP, MTR2, pDM and unique ENL networks, we used a genetic matching algorithm[55] to balance HC and SZ cohorts for confounding factors and we subsequently analyzed networks derived from the balanced cohorts (detailed methods are provided in Supplementary Note 2). The matched analysis revealed a greater number of significant ENL voxels relative to LIN (Supplementary Fig. 3 and Supplementary Table 4) and validated our primary findings, strongly indicating that ENL estimates of the networks in question outperform LIN in capturing SZ FC alterations (summary test information is provided in Supplementary Table 4).

Discussion

Linear FC analysis remains a fruitful method for extracting valuable information from fMRI data. However, despite its usefulness and ease of interpretation, various brain processes exhibit nonlinear aspects[15,17], suggesting that linear FC provides us with a limited view of the data and the clinical neurocognitive hypothesis space. Previous rsfMRI studies have identified evidence of nonlinearity and its prospective role in differentiating cohorts[37,41,42], but our approach to ENL ICN estimation demonstrates the potential of connectivity-domain ICA[33] and nonlinear information to shape the predictive clinical landscape and inform systems neuroscience theorizing.

We find that components extracted from ENL-wFC exhibit higher reliability than those extracted from LIN-wFC, and that unique networks are identified from each FC estimator. Our validation analysis supports these findings. We also find that corresponding networks exhibit striking spatial variation. Among potential explanations, the presence of greater ENL weight within core regions could be reflective of stronger signals within core areas. However, we note that such a hypothesis probably cannot explain the detection of differences between the HC and SZ cohorts. Because ENL ICNs represent independent data sources composed of elements whose distance correlation values deviate from a linear relationship with Pearson correlation, the identified gradients may reflect actual differences in the underlying FC complexity, which merits further investigation of their potential cognitive and clinical significance. Notably, higher ENL weight within core regions indicates that linear connectivity analyses may underestimate FC within network centers. Future work will investigate potential explanations for the observed gradients.

The recovery of a unique ENL network underscores the importance of effectively capturing networks that accurately reflect nonlinear connectivity information, as our results show that networks estimated from information not explained by linear connectivity may be altered in psychiatric conditions such as SZ. For instance, the unique ENL network consolidated regions typically associated with cingulo-opercular[56], anterior default mode[57] and central executive[58] networks, suggesting that this ICN may act as an integrator hub for multiple large-scale brain ensembles. This hypothesis is consistent with mounting evidence of a role for anterior insular regions in mediating information flow between default mode and central executive regions[59–61]. Moreover, anterior insular regions are associated with event and stimulus salience processing, both of which are reported to be compromised in SZ[62]. Importantly, group comparisons revealed functional hypoconnectivity for SZ within the left anterior insula of the unique ENL ICN, suggesting that our method can capture hidden patterns that reflect inefficiencies in the integration of brain networks in psychosis. This finding serves as a case in point connecting our methodology to the generation of novel insights, and demonstrates the potential of our approach to contribute to the development of brain-based biomarkers of psychiatric disorders.

The finding that ENL voxels collectively exhibit a greater degree of sensitivity to SZ diagnosis further elaborates the potential of nonlinear connectivity information to play a role within clinical FC analysis. ENL TEMP comparisons revealed larger clusters of significant voxels within auditory and language-related regions that have been previously associated with SZ and positive symptoms such as auditory verbal hallucinations in both tfMRI[63,64] and rsfMRI[65,66] analyses. For example, ENL tests revealed expansive clusters within superior temporal regions known to implement acoustic–phonetic computations[52]. ENL tests also identified a sizable volume with higher SZ values within the right supramarginal gyrus, which has been shown to play a role in phonological decision-making[67]. By contrast, the right supramarginal gyrus was almost entirely missed for significance by LIN TEMP comparisons. Notably, ENL MTR2 comparisons revealed greater numbers of significant voxels within sensorimotor regions previously implicated in SZ[66,68], as well as clusters within the bilateral posterior insula that were not detected by LIN tests. Moreover, ENL pDM comparisons revealed hyperconnectivity for SZ that was missed by LIN within core regions of the pDM that have been associated with reflective, internally focused cognitive processes thought to be relevant to SZ diagnosis and symptoms[69]. This finding was validated by the matched cohort analysis (Supplementary Fig. 3). Overall, our results demonstrate that nonlinear statistical dependencies in fMRI data can be leveraged to distinguish these cohorts and warrant further investigation of the relationship between features extracted from measures that are sensitive to nonlinearity and the presentation of psychosis.

Our previous work proposed this conceptual framework[39]. Here, we advance and rigorously investigate the framework by providing an in-depth quantitative analysis of ENL and LIN networks, their spatial variation and their sensitivity to differences between HC and SZ. However, the current analysis has several methodological and interpretive limitations. First, although we utilized a large dataset collected across multiple psychosis projects and sites to address representative sampling issues[45], the generalizability of our results is limited to populations of individuals with demographic characteristics similar to that of the analyzed sample. For instance, reporting on race for the present study utilized three super-population groupings (mixed American, European and African), and our results cannot necessarily be generalized to populations that fall outside these groups. Second, we note that alternate models of the relationship between NL-wFC and LIN-wFC can be leveraged when estimating ENL-wFC. Therefore, we do not claim that the current method of estimation is decisive or definitive to the potential exclusion of methods designed to estimate ENL-wFC using alternate models. Future work will investigate the use of other models with the aim of providing increasingly robust and precise characterizations of whole-brain connectivity features not explained by linear connectivity patterns. Third, we note that although our approach may share conceptual similarities with methods that construct nonlinear fMRI connectivity using features derived from pairwise associations[41,42], we do not necessarily expect the findings of these distinct approaches to converge due to substantial differences in methodology. Thus, we leave any speculation about the relationship between features extracted from these methods as an open empirical question for future investigation. Fourth, we note that attributing context-invariant functions to macroscopic brain networks may oversimplify their roles. The functions attributed in the present study are suggested as those among the most supported by previous research findings. Finally, although our results warrant further investigation into the potential neurocognitive and psychiatric roles of ENL networks, we maintain that moving beyond association will probably require developing interventions that can effectively tie the extracted features to the causal outcomes of cognitive operations, psychiatric diagnosis and symptoms.

The primary goal of the current study was to investigate the presence and clinical utility of nonlinear FC patterns that move beyond linear FC. However, future work will investigate networks extracted from nonlinear FC patterns in the context of task-based experimental designs. Additionally, future work will focus on replicating our results in large-scale B-SNIP transdiagnostic rsfMRI datasets[70,71], on utilizing ENL networks to distinguish a broader array of clinical cohorts, on analyzing associations with cognitive and symptom scores, and on analyzing the temporal[72] and spatial[32,40,66,73,74] dynamics exhibited by NL and ENL networks during task performance and at rest.

Methods

Subject information, data acquisition and quality control

We analyzed 3-Tesla rsfMRI data sourced from three case–control psychosis projects—COBRE, FBIRN and MPRC (Fig. 2 and Table 1). Detailed subject recruitment information, as well as inclusion and exclusion criteria for COBRE, FBIRN and MPRC studies, can respectively be found in refs. 43,44,46 as well as in the Reporting Summary. Sex was based on self-reported demographic assessment. Race was based on a multi-dimensional scaling (MDS) analysis conducted on combined local samples and data from the the 1000 Genomes Project[75]. For each super-population of mixed American (AMR), European (EUR) and African (AFR) individuals, a cluster centroid was obtained based on 1000 Genomes data. Local samples were assigned to the nearest reference population, and those that were distant (>3 s.d. away) from any population cluster were assigned to the 'other' category. Subjects provided informed written consent as required and approved by the Institutional Review Boards (IRBs) of the corresponding institutions. COBRE participants gave written informed consent as required and approved by the IRB of the University of New Mexico[43]. FBIRN participants gave written informed consent as required and approved by the IRBs of the University of California Irvine, the University of California Los Angeles, the University of California San Francisco, Duke University, University of North Carolina, University of New Mexico, University of Iowa and University of Minnesota[44]. MPRC participants gave written informed consent as required and approved by the IRB of the University of Maryland, Baltimore[46].

Individuals with SZ from the COBRE dataset received a diagnosis of schizophrenia performed by two research psychiatrists in consensus via the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) using the patient version of the SCID-DSM-IV-TR[43]. SZ subjects were evaluated for comorbidities and for retrospective as well as prospective clinical stability. Individuals with SZ from the FBIRN study were diagnosed with schizophrenia based on the SCID-DSM-IV-TR and were clinically stable for at least two months before scanning[44]. For MPRC SZ subjects, a diagnosis of schizophrenia was confirmed via the SCID-DSM-IV[46]. Case and control participants were compensated for interviews, scan sessions and assessments conducted during the referenced studies.

COBRE data were collected at a single site on a Siemens TIM Trio scanner via an echo-planar imaging sequence (repetition time (TR) = 2,000 ms; echo time (TE) = 29 ms)[20]. Voxel spacing was 3.75 × 3.75 × 4.5 mm, the slice gap was 1.05 mm, and the field of view (FOV) was 240 × 240 mm. FBIRN data were collected from seven sites[76], with six sites utilizing Siemens TIM Trio scanners and one utilizing a General Electric Discovery MR750 system[20]. All seven sites used an echo-planar imaging sequence (TR = 2,000 ms; TE = 30 ms). The original voxel spacing was 3.4375 × 3.4375 × 4 mm, the slice gap was 1 mm, and the FOV was 220 × 220 mm. MPRC data were collected from three sites via echo-planar imaging sequences[20]. One site used a Siemens Allegra scanner (TR = 2,000 ms; TE = 27 ms; voxel spacing = 3.44 × 3.44 × 4 mm; FOV = 220 × 220 mm), another used a Siemens TIM Trio scanner (TR = 2,210 ms; TE = 30 ms; voxel spacing = 3.44 × 3.44 × 4 mm; FOV = 220 × 220 mm), and the third site used a Siemens TIM Trio scanner (TR = 2,000 ms; TE = 30 ms; voxel spacing = 1.72 × 1.72 × 4 mm; FOV = 220 × 220 mm).

The following subject quality control criteria[21] were used for the current study: (1) completeness of demographic information, (2) availability of T1 structural MRI, (3) availability of genomic information, (4) maximum head rotation less than 3°, (5) maximum translation less than 3 mm, (6) mean framewise displacement less than 0.25, (7) quality registration to an echo-planar imaging template, (8) whole-brain (in addition to the top ten and bottom ten slices) spatial overlap between the subject mask and group mask greater than 80% and (9) removal of duplicate subjects. The final subject pool included 315 HC and 193 SZ ($n = 508$) individuals.

Preprocessing

Preprocessing was performed primarily within the MATLAB software environment using Statistical Parametric Mapping (SPM 12; http://www.fil.ion.ucl.ac.uk/spm/) and the FMRIB Software Library (FSL v6.0; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki). Preprocessing steps included (1) rigid body motion and slice timing correction, (2) nonlinear warping to the Montreal Neurological Institute 152 coordinate space, (3) spatial resampling to 3-mm isotropic voxel spacing, (4) spatial smoothing with a 6-mm full-width at half-maximum Gaussian kernel, (5) head motion regression, detrending, despiking and low-pass filtering, (6) temporal resampling to TR = 2,000 ms and (7) voxel time-series Z-scoring to normalize the variance.

Constructing LIN and ENL FC

We constructed LIN as well as ENL global (voxel-wise) FC matrices for every subject[39]. Let $X \in \mathbb{R}^{n \times v}$ be a sample of rsfMRI data where n is the number of time points, v is the number of voxels within the brain, and x and y represent any two preprocessed voxel time series such that $x, y \in \mathbb{R}^{1 \times n}$. Thus, x_i is the value of voxel x at time point i. We estimated each subject's LIN-wFC as the covariance (Cov) across all pairs of brain voxels (equation (1)). Because voxel time courses were Z-scored during preprocessing, the pairwise covariance was equal to the pairwise Pearson correlation, which was used conventionally to estimate linear FC:

$$\text{LIN}_{\text{wFC}_{x,y}} = \text{Cov}(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i)(y_i) \tag{1}$$

Next, we calculated the voxel-wise distance correlation[38] to construct NL-wFC. Distance correlation is a representation of the association between random vectors based on Euclidean distances between sample observations[38] (equation (2)):

$$\text{NL}_{\text{wFC}_{x,y}} = \text{dCorr}(x, y) = \frac{\text{dCov}(x, y)}{\sqrt{\text{dVar}(x)\,\text{dVar}(y)}} \tag{2}$$

where

$$\text{dCov}_n^2(x, y) = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{n} A_{j,k} B_{j,k}$$

and

$$\text{dVar}_n^2(x) = \text{dCov}_n^2(x, x) = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} A_{k,l}^2$$

The squared sample distance covariance (dCov²) is calculated as the arithmetic average of products AB, where A and B represent the doubly centered Euclidean distance matrices of rsfMRI voxel time series x and y such that

$$a_{j,k} = \|x_j - x_k\| \quad j, k = 1, 2, \dots, n$$

$$b_{j,k} = \|y_j - y_k\|, \quad j, k = 1, 2, \dots, n$$

$$A_{j,k} = a_{j,k} - \bar{a}_{j\cdot} - \bar{a}_{\cdot k} - \bar{a}_{\cdot\cdot},$$

$$B_{j,k} = b_{j,k} - \bar{b}_{j\cdot} - \bar{b}_{\cdot k} - \bar{b}_{\cdot\cdot}$$

We note that distance correlation is sensitive to both linear and nonlinear dependence relations, and that the distance correlation between random vectors is zero if and only if the vectors are independent[38].

Because we are interested in extracting networks from distance correlation patterns that are not explained by Pearson correlation, we removed the effect of LIN-wFC on NL-wFC using an ordinary least-squares approach to estimate the ENL-wFC for each subject (equation (3)). We first vectorized both NL-wFC and LIN-wFC. We then removed the linear relationship between NL-wFC and LIN-wFC using a regression-based method and reshaped the vector of residuals into a $v \times v$ FC matrix:

$$\text{ENL}_{\text{wFC}} = \text{vec}^{-1}(\text{vec}(\text{NL}_{\text{wFC}}) - \alpha \times \text{vec}(\text{LIN}_{\text{wFC}})) \tag{3}$$

where

$$\min_{\alpha} \sum_{i=1}^{v^2} ((\text{vec}(\text{NL}_{\text{wFC}}))_i - (\text{vec}(\text{LIN}_{\text{wFC}}))_i)^2$$

We treated the estimation of α as an ordinary least-squares problem by finding the value of α that minimized the sum of squared errors between NL-wFC and LIN-wFC. Thus, here we define the ENL-wFC for a given subject as the NL-wFC information with the linear effect of LIN-wFC removed. For each subject, the goodness of fit of the linear model was evaluated via the coefficient of determination (R^2). To assess the difference in R^2 between HC and SZ cohorts, we used a general linear model (GLM) to remove the effect of confounding factors commonly reported in psychosis studies, including age, sex, site and motion (mean framewise displacement), on the goodness-of-fit data, and we subsequently conducted a two-sided permutation test with 5,000 random permutations (Krol, 2023; https://github.com/lrkrol/permutationTest)[77].

Extracting ICNs

We used the Group ICA of the fMRI Toolbox (GIFT v4.0; http://trendscenter.org/software/gift)[72] to implement connectivity-domain ICA[33] and obtain separate sets of group-level networks from the LIN-wFC and ENL-wFC data. The implementation of gr-sICA was preceded by an initial subject-level, multi-power iteration[78], principal component analysis step to reduce dimensionality and denoise the data[79]. The 30 principal components that explained the maximum variance of each subject's respective LIN-wFC and ENL-wFC were retained for further analysis. Subject-level principal components from each estimator were concatenated across the component dimension, and a group-level principal component analysis step was applied to further reduce the dimensionality of the data and decrease the computational demands of gr-sICA[11]. The 20 group-level principal components that explained the maximum variance of each estimator-specific dataset were used as the input for gr-sICA. We selected a gr-sICA model order of 20 to obtain large-scale functional networks[33,80]. To ensure the reliability of our results, ICA was implemented via the Infomax optimization algorithm[81] 100 times, with both random initialization and bootstrapping, and the most stable run was selected for further analysis. We evaluated the reliability and quality of ENL and LIN components using the ICASSO quality index (IQ), which quantifies component stability across runs[82]. To assess the difference in stability between ENL and LIN components, we conducted a two-sided permutation test with 5,000 random permutations on the IQ data. Assessing component reliability was a necessary step, as previous work has demonstrated that certain components may be inconsistently extracted from the data of interest[82]. In the context

of fMRI network estimation, ICASSO IQ is often used to differentiate reliable components from components that are unstable and unfit for further analysis[66]. A component was identified as an ICN if and only if (1) it exhibited an ICASSO IQ value exceeding 0.80, (2) it exhibited high visual overlap with gray matter, (3) it exhibited peak weight within gray matter, and (4) it exhibited low visual similarity to motion, ventricular and other known artefacts. To find corresponding networks, the spatial correlation value was computed between every pair of extracted LIN and ENL components, and components were matched in a greedy fashion. ICNs matched with a spatial correlation value exceeding 0.80 were classified as common[21] and were labeled based on their neuroanatomical distributions and the identification of ICNs from previous studies[33]. Networks exhibiting a maximum spatial correlation of less than 0.40 were classified as unique. We used the Group ICA of fMRI Toolbox (GIFT v4.0) to implement group information-guided ICA (GIG-ICA)[83] and reconstruct subject-specific networks from subject-level principal components using the group-level spatial references.

Assessment of spatial variation among corresponding ICNs

To assess differences in spatial variation between matched networks, we conducted voxel-wise, two-sided, paired-samples t-tests on their Z-scored subject-level estimates. For a given matched network pair, statistical comparisons were masked for voxels exceeding $Z = 1.96$ ($P = 0.05$) in either group-level map (LIN or ENL), and the false discovery rate (FDR)[84] method was used to correct for multiple comparisons ($q < 0.05$). The robustness of the voxel-wise t-test procedure was assessed via comparison to the results of voxel-wise two-sided permutation tests with 5,000 random permutations for the pDM network. The automated anatomical labeling atlas 3 (AAL3)[85] was used to localize clusters of significant voxels to anatomically defined brain regions.

Assessment of ICN differences between HC and SZ

To assess the differences between HC and SZ, we conducted voxel-wise, independent-samples t-tests between the estimates of common and unique networks derived from each cohort. We first used a GLM to remove the effects of confounding factors such as age, sex, site and motion (mean framewise displacement) on Z-scored subject-level network estimates. Voxel-wise, two-sided, independent-samples t-tests were then conducted on the residual spatial maps derived from the HC and SZ groups. Statistical comparisons between common networks were masked for voxels exceeding $Z = 1.96$ ($P = 0.05$) in either of the group-level maps (LIN or ENL), and unique network comparisons were masked for voxels exceeding the same threshold in the unique group-level map. The FDR[84] method was used to correct for multiple comparisons ($q < 0.05$). For both ENL and LIN, the robustness of the voxel-wise t-test procedure was assessed via comparison to the results of voxel-wise, two-sided permutation tests with 5,000 random permutations for the pDM network. The AAL3[85] atlas was used to localize clusters of significant voxels to anatomically defined brain regions. A two-sided McNemar's test was used to assess the overall ENL versus LIN difference in statistical sensitivity (across all voxels belonging to commonly classified networks), and differences in statistical sensitivity for matched network pairs were investigated separately using either two-sided McNemar's tests or exact binomial tests (for $n < 25$).

Reporting summary

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

Data availability

Contact information and resources for obtaining further details for the private datasets utilized in the present study are as follows. COBRE: Vince D. Calhoun (vcalhoun@gsu.edu), Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA[43]. FBIRN: Theo G. M. van Erp (tvanerp@hs.uci.edu),

Clinical Translational Neuroscience Laboratory, Department of Psychiatry and Human Behavior, University of California, Irvine, CA, USA[45]. MPRC: Peter Kochunov (ms.psychiatry@uth.tmc.edu), Department of Psychiatry and Behavioral Science, University of Texas Health Science Center Houston, Houston, TX[46].

Code availability

Preprocessing and data analysis were conducted primarily within the MATLAB software environment mainly using MATLAB 9.9.0.1857802 (R2020b) Update 7, the Statistical Parametric Mapping toolbox (SPM 12), the FMRIB software library (FSL v6.0), the Group ICA of fMRI toolbox (GIFT v4.0) and RStudio (R v4.1.2). MATLAB R2020b can be downloaded from https://www.mathworks.com. The FSL v6.0 toolbox can be downloaded from https://fsl.fmrib.ox.ac.uk/fsl/fslwiki. The SPM12 toolbox can be downloaded from https://www.fil.ion.ucl.ac.uk/spm/. GIFT v4.0 can be downloaded from https://trendscenter.org/software/gift/. R v4.1.2 can be downloaded from https://cran.r-project.org/. The sample scripts utilized for dual code data visualization[86] can be downloaded from https://trendscenter.org/x/datavis/. The permutation test function utilized for statistical randomization analyses (Krol, 2023)[77] can be downloaded from https://github.com/lrkrol/permutationTest/. The R package used to balance HC and SZ cohorts for confounding factors and assess differences using Abadie–Imbens standard errors[87] can be downloaded from https://CRAN.R-project.org/package=Matching. The MATLAB function used to calculate ENL-wFC can be downloaded from https://github.com/trendscenter/calc_ENLwFC. Other MATLAB code used for this study can be obtained from the corresponding authors upon reasonable request.

References

1. Voineskos, A. N. et al. Functional magnetic resonance imaging in schizophrenia: current evidence, methodological advances, limitations and future directions. *World Psychiatry* **23**, 26–51 (2024).
2. Logothetis, N. K. The underpinnings of the BOLD functional magnetic resonance imaging signal. *J. Neurosci.* **23**, 3963–3971 (2003).
3. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
4. Magri, C., Schridde, U., Murayama, Y., Panzeri, S. & Logothetis, N. K. The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. *J. Neurosci.* **32**, 1395–1407 (2012).
5. Pan, W. J., Thompson, G. J., Magnuson, M. E., Jaeger, D. & Keilholz, S. Infraslow LFP correlates to resting-state fMRI BOLD signals. *NeuroImage* **74**, 288–297 (2013).
6. Shi, Z. et al. High spatial correspondence at a columnar level between activation and resting state fMRI signals and local field potentials. *Proc. Natl Acad. Sci. USA* **114**, 5253–5258 (2017).
7. Dong, D., Wang, Y., Chang, X., Luo, C. & Yao, D. Dysfunction of large-scale brain networks in schizophrenia: a meta-analysis of resting-state functional connectivity. *Schizophr. Bull.* **44**, 168–181 (2018).
8. Li, S. et al. Dysconnectivity of multiple brain networks in schizophrenia: a meta-analysis of resting-state functional connectivity. *Front. Psychiatry* **10**, 482 (2019).
9. Sheffield, J. M. & Barch, D. M. Cognition and resting-state functional connectivity in schizophrenia. *Neurosci. Biobehav. Rev.* **61**, 108–120 (2016).
10. Friston, K., Brown, H. R., Siemerkus, J. & Stephan, K. E. The dysconnection hypothesis (2016). *Schizophr. Res.* **176**, 83–94 (2016).
11. Calhoun, V. D., Liu, J. & Adali, T. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic and ERP data. *NeuroImage* **45**, S163–S172 (2009).

12. Friston, K. J. Functional and effective connectivity: a review. *Brain Connect.* **1**, 13–36 (2011).

13. Mohanty, R., Sethares, W. A., Nair, V. A. & Prabhakaran, V. Rethinking measures of functional connectivity via feature extraction. *Sci. Rep.* **10**, 1298 (2020).

14. Deshpande, G., LaConte, S., Peltier, S. & Hu, X. Connectivity analysis of human functional MRI data: from linear to nonlinear and static to dynamic. In *Proc. Medical Imaging and Augmented Reality*: *Third International Workshop* 17–24 (Springer, 2006).

15. Friston, K. J. Brain function, nonlinear coupling and neuronal transients. *Neuroscientist* **7**, 406–418 (2001).

16. He, F. & Yang, Y. Nonlinear system identification of neural systems from neurophysiological signals. *Neuroscience* **458**, 213–228 (2021).

17. Singer, W. Cortical dynamics revisited. *Trends Cogn. Sci.* **17**, 616–626 (2013).

18. Stephan, K. E. et al. Nonlinear dynamic causal models for fMRI. *NeuroImage* **42**, 649–662 (2008).

19. Korhonen, O., Zanin, M. & Papo, D. Principles and open questions in functional brain network reconstruction. *Hum. Brain Mapp.* **42**, 3680–3711 (2021).

20. Iraji, A. et al. Multi-spatial-scale dynamic interactions between functional sources reveal sex-specific changes in schizophrenia. *Netw. Neurosci.* **6**, 357–381 (2022).

21. Iraji, A. et al. Identifying canonical and replicable multi-scale intrinsic connectivity networks in 100k+ resting-state fMRI datasets. *Hum. Brain Mapp.* **44**, 5729–5748 (2023).

22. Shamir, M. & Sompolinsky, H. Nonlinear population codes. *Neural Comput.* **16**, 1105–1136 (2004).

23. Rashid, B. & Calhoun, V. Towards a brain-based predictome of mental illness. *Hum. Brain Mapp.* **41**, 3468–3535 (2020).

24. Adali, T., Anderson, M. & Fu, G. S. Diversity in independent component and vector analyses: identifiability, algorithms and applications in medical imaging. *IEEE Signal Process. Mag.* **31**, 18–33 (2014).

25. Comon, P. & Jutten, C. *Handbook of Blind Source Separation*: *Independent Component Analysis and Applications* (Academic, 2010).

26. Calhoun, V. D., Kiehl, K. A. & Pearlson, G. D. Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. *Hum. Brain Mapp.* **29**, 828–838 (2008).

27. Iraji, A. et al. Moving beyond the 'CAP' of the iceberg: intrinsic connectivity networks in fMRI are continuously engaging and overlapping. *NeuroImage* **251**, 119013 (2022).

28. Seeley, W. W. et al. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007).

29. Damoiseaux, J. S. et al. Consistent resting-state networks across healthy subjects. *Proc. Natl Acad. Sci. USA* **103**, 13848–13853 (2006).

30. Laird, A. R. et al. Behavioral interpretations of intrinsic connectivity networks. *J. Cogn. Neurosci.* **23**, 4022–4037 (2011).

31. Wu, L., Caprihan, A. & Calhoun, V. Tracking spatial dynamics of functional connectivity during a task. *NeuroImage* **239**, 118310 (2021).

32. Iraji, A. et al. Spatial dynamics within and between brain functional domains: a hierarchical approach to study time-varying brain function. *Hum. Brain Mapp.* **40**, 1969–1986 (2019).

33. Iraji, A. et al. The connectivity domain: analyzing resting state fMRI data using feature-based data-driven and model-based methods. *NeuroImage* **134**, 494–507 (2016).

34. Wu, L., Caprihan, A., Bustillo, J., Mayer, A. & Calhoun, V. An approach to directly link ICA and seed-based functional connectivity: application to schizophrenia. *NeuroImage* **179**, 448–470 (2018).

35. Calhoun, V. D. & Allen, E. Extracting intrinsic functional networks with feature-based group independent component analysis. *Psychometrika* **78**, 243–259 (2013).

36. Hyvärinen, A., Sasaki, H. & Turner, R. Nonlinear ICA using auxiliary variables and generalized contrastive learning. In *Proc. 22nd International Conference on Artificial Intelligence and Statistics* Vol. 89, 859–868 (PMLR, 2019).

37. Morioka, H., Calhoun, V. & Hyvärinen, A. Nonlinear ICA of fMRI reveals primitive temporal structures linked to rest, task, and behavioral traits. *NeuroImage* **218**, 116989 (2020).

38. Székely, G. J., Rizzo, M. L. & Bakirov, N. K. Measuring and testing dependence by correlation of distances. *Ann. Stat.* **35**, 2769–2794 (2007).

39. Iraji, A. et al. The nonlinear brain: towards uncovering hidden brain networks using explicitly nonlinear functional interaction. In *Proc. IEEE International Symposium on Biomedical Imaging* 1–4 (IEEE, 2023); https://doi.org/10.1109/ISBI53787.2023.10230347

40. Bhinge, S., Long, Q., Calhoun, V. D. & Adali, T. Spatial dynamic functional connectivity analysis identifies distinctive biomarkers in schizophrenia. *Front. Neurosci.* **13**, 1006 (2019).

41. Motlaghian, S. M. et al. Nonlinear functional network connectivity in resting functional magnetic resonance imaging data. *Hum. Brain Mapp.* **43**, 4556–4566 (2022).

42. Motlaghian, S. M. et al. A method for estimating and characterizing explicitly nonlinear dynamic functional network connectivity in resting-state fMRI data. *J. Neurosci. Methods* **389**, 109794 (2023).

43. Aine, C. J. et al. Multimodal neuroimaging in schizophrenia: description and dissemination. *Neuroinformatics* **15**, 343–364 (2017).

44. Damaraju, E. et al. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage Clin.* **5**, 298–308 (2014).

45. Keator, D. B. et al. The Function Biomedical Informatics Research Network data repository. *NeuroImage* **124**, 1074–1079 (2016).

46. Adhikari, B. M. et al. Functional network connectivity impairments and core cognitive deficits in schizophrenia. *Hum. Brain Mapp.* **40**, 4593–4605 (2019).

47. Smith, S. M. et al. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl Acad. Sci. USA* **106**, 13040–13045 (2009).

48. Moerel, M., De Martino, F. & Formisano, E. An anatomical and functional topography of human auditory cortical areas. *Front. Neurosci.* **8**, 225 (2014).

49. Szczepanski, S. M., Pinsk, M. A., Douglas, M. M., Kastner, S. & Saalmann, Y. B. Functional and structural architecture of the human dorsal frontoparietal attention network. *Proc. Natl Acad. Sci. USA* **110**, 15806–15811 (2013).

50. Wang, S., Tepfer, L. J., Taren, A. A. & Smith, D. V. Functional parcellation of the default mode network: a large-scale meta-analysis. *Sci. Rep.* **10**, 16096 (2020).

51. Niendam, T. A. et al. Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* **12**, 241–268 (2012).

52. Bhaya-Grossman, I. & Chang, E. F. Speech computations of the human superior temporal gyrus. *Annu. Rev. Psychol.* **73**, 79–102 (2022).

53. Rupp, K. et al. Neural responses in human superior temporal cortex support coding of voice representations. *PLoS Biol.* **20**, e3001675 (2022).

54. Caspers, J. et al. Within- and across-network alterations of the sensorimotor network in Parkinson's disease. *Neuroradiology* **63**, 2073–2085 (2021).

55. Sekhon, J. S. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. *J. Stat. Software* **42**, 1–52 (2011).

56. Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L. & Petersen, S. E. A dual-networks architecture of top-down control. *Trends Cogn. Sci.* **12**, 99–105 (2008).

57. Sendi, M. S. E. et al. Aberrant dynamic functional connectivity of default mode network in schizophrenia and links to symptom severity. *Front. Neural Circuits* **15**, 649417 (2021).

58. Heinonen, J. et al. Default mode and executive networks areas: association with the serial order in divergent thinking. *PLoS ONE* **11**, e0162234 (2016).

59. Molnar-Szakacs, I. & Uddin, L. Q. Anterior insula as a gatekeeper of executive control. *Neurosci. Biobehav. Rev.* **139**, 104736 (2022).

60. Uddin, L. Q., Nomi, J. S., Hébert-Seropian, B., Ghaziri, J. & Boucher, O. Structure and function of the human insula. *J. Clin. Neurophysiol.* **34**, 300–306 (2017).

61. Uddin, L. Q., Yeo, B. T. T. & Spreng, R. N. Towards a universal taxonomy of macro-scale functional human brain networks. *Brain Topogr.* **32**, 926–942 (2019).

62. Palaniyappan, L. & Liddle, P. F. Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *J. Psychiatry Neurosci.* **37**, 17–27 (2012).

63. Calhoun, V. D. et al. Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder. *Front. Psychiatry* **2**, 75 (2012).

64. Kim, D. I. et al. Auditory oddball deficits in schizophrenia: an independent component analysis of the fMRI multisite function BIRN study. *Schizophr. Bull.* **35**, 67–81 (2009).

65. Alderson-Day, B., McCarthy-Jones, S. & Fernyhough, C. Hearing voices in the resting brain: a review of intrinsic functional connectivity research on auditory verbal hallucinations. *Neurosci. Biobehav. Rev.* **55**, 78–87 (2015).

66. Iraji, A. et al. The spatial chronnectome reveals a dynamic interplay between functional segregation and integration. *Hum. Brain Mapp.* **40**, 3058–3077 (2019).

67. Hartwigsen, G. et al. Phonological decisions require both the left and right supramarginal gyri. *Proc. Natl Acad. Sci. USA* **107**, 16494–16499 (2010).

68. Kaufmann, T. et al. Disintegration of sensorimotor brain networks in schizophrenia. *Schizophr. Bull.* **41**, 1326–1335 (2015).

69. Garrity, A. G. et al. Aberrant 'default mode' functional connectivity in schizophrenia. *Am. J. Psychiatry* **164**, 450–457 (2007).

70. Meda, S. A. et al. Differences in resting-state functional magnetic resonance imaging functional network connectivity between schizophrenia and psychotic bipolar probands and their unaffected first-degree relatives. *Biol. Psychiatry* **71**, 881–889 (2012).

71. Meda, S. A. et al. Frequency-specific neural signatures of spontaneous low-frequency resting state fluctuations in psychosis: evidence from Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) Consortium. *Schizophr. Bull.* **41**, 1336–1348 (2015).

72. Iraji, A. et al. Tools of the trade: estimating time-varying connectivity patterns from fMRI data. *Soc. Cogn. Affect. Neurosci.* **16**, 849–874 (2021).

73. Iraji, A., Miller, R., Adali, T. & Calhoun, V. D. Space: a missing piece of the dynamic puzzle. *Trends Cogn. Sci.* **24**, 135–149 (2020).

74. Long, Q., Bhinge, S., Calhoun, V. D. & Adali, T. Graph-theoretical analysis identifies transient spatial states of resting-state dynamic functional network connectivity and reveals dysconnectivity in schizophrenia. *J. Neurosci. Methods* **350**, 109039 (2021).

75. 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

76. Turner, J. A. et al. A multi-site resting state fMRI study on the amplitude of low frequency fluctuations in schizophrenia. *Front. Neurosci.* **7**, 137 (2013).

77. Krol permutationTest. GitHub https://github.com/lrkrol/permutationTest (2023).

78. Rachakonda, S., Silva, R. F., Liu, J. & Calhoun, V. D. Memory efficient PCA methods for large group ICA. *Front. Neurosci.* **10**, 17 (2016).

79. Erhardt, E. B. et al. Comparison of multi-subject ICA methods for analysis of fMRI data. *Hum. Brain Mapp.* **32**, 2075–2095 (2011).

80. Ray, K. L. et al. ICA model order selection of task co-activation networks. *Front. Neurosci.* **7**, 237 (2013).

81. Bell, A. J. & Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995).

82. Himberg, J., Hyvärinen, A. & Esposito, F. Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage* **22**, 1214–1222 (2004).

83. Du, Y. & Fan, Y. Group information guided ICA for fMRI data analysis. *NeuroImage* **69**, 157–197 (2013).

84. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B (Methodol.)* **57**, 289–300 (1995).

85. Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J. & Joliot, M. Automated anatomical labelling atlas 3. *NeuroImage* **206**, 116189 (2020).

86. Allen, E. A., Erhardt, E. B. & Calhoun, V. D. Data visualization in the neurosciences: overcoming the curse of dimensionality. *Neuron* **74**, 603–608 (2012).

87. Abadie, A. & Imbens, G. W. Large sample properties of matching estimators for average treatment effects. *Econometrica* **74**, 235–267 (2006).

Acknowledgements

Author contributions

A.I. and V.D.C. proposed the study. V.D.C., P.K., B.M.A. and T.G.M.v.E. acquired the original data. S.K., A.I., V.D.C., J.C. and K.K. contributed to methods development. S.K. analyzed the data. S.K., A.I. and V.D.C. contributed to the interpretation of the results. S.K. drafted the paper. A.I., V.D.C., P.A.C., J.C., T.A., J.F., T.G.M.v.E. and M.D. contributed to the critical revision of the paper.

Competing interests

The authors declare no competing interests.

Ethics Statement

This observational study was conducted retrospectively using data acquired from human subjects in compliance with all relevant ethical regulations.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44220-024-00341-y.

Correspondence and requests for materials should be addressed to Spencer Kinsey or Armin Iraji.

Peer review information *Nature Mental Health* thanks Elvisha Dhamala, Xin Di and the other, anonymous, reviewers for their contribution to the peer review of this work.

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

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

[1]Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA. [2]Neuroscience Institute, Georgia State University, Atlanta, GA, USA. [3]Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway. [4]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, liSGM, CIBERSAM, School of Medicine, Universidad Complutense, Madrid, Spain. [5]Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD, USA. [6]Department of Psychiatry and Behavioral Science, University of Texas Health Science Center at Houston, Houston, TX, USA. [7]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA, USA. [8]San Francisco Veterans Affairs Medical Center, San Francisco, CA, USA. [9]Clinical Translational Neuroscience Laboratory, Department of Psychiatry and Human Behavior, University of California, Irvine, CA, USA. [10]Department of Physics and Astronomy, Georgia State University, Atlanta, GA, USA. [11]Department of Computer Science, Georgia State University, Atlanta, GA, USA. ✉e-mail: skinsey8@gsu.edu; airaji@gsu.edu

nature portfolio

Corresponding author(s): Spencer Kinsey and Armin Iraji

Last updated by author(s): 08/02/2024

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	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	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)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	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. The present study was based on referenced datasets.
Data analysis	Preprocessing and data analysis were conducted primarily within the MATLAB software environment mainly using MATLAB 9.9.0.1857802 (R2020b) Update 7, the Statistical Parametric Mapping toolbox (SPM 12), the FMRIB software library (FSL v6.0), the Group ICA of fMRI toolbox (GIFT v4.0), and RStudio (R v4.1.2). MATLAB R2020b can be downloaded from https://www.mathworks.com. The FSL v6.0 toolbox can be downloaded from https://fsl.fmrib.ox.ac.uk/fsl/fslwiki. The SPM 12 toolbox can be downloaded from https://www.fil.ion.ucl.ac.uk/spm/. GIFT v4.0 can be downloaded from https://trendscenter.org/software/gift/. R v4.1.2 can be downloaded from https://cran.r-project.org/. The sample scripts utilized for dual code data visualization (Allen et al., 2012; https://doi.org/10.1016/j.neuron.2012.05.001) can be downloaded from https://trendscenter.org/x/datavis/. The permutation test function utilized for statistical randomization analyses (Krol, 2023) can be downloaded from https://github.com/lrkrol/permutationTest/.

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.

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

Contact information and resources for obtaining further details for the private datasets utilized in the present study are listed as follows:

COBRE: Vince D. Calhoun (vcalhoun@gsu.edu), Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA
(Aine et al., 2017; https://doi.org/10.1007/s12021-017-9338-9)

FBIRN: Theo G. M. van Erp (tvanerp@hs.uci.edu), Clinical Translational Neuroscience Laboratory, Department of Psychiatry and Human Behavior, University of California, Irvine, CA, USA
(Keator et al., 2016; https://doi.org/10.1016/j.neuroimage.2015.09.003)

MPRC: Peter Kochunov (ms.psychiatry@uth.tmc.edu), Department of Psychiatry and Behavioral Science, University of Texas Health Science Center Houston, Houston, TX
(Adhikari et al., 2019; https://doi.org/10.1002/hbm.24723)

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Our findings apply to both male and female sexes. For each referenced dataset (COBRE, FBIRN, and MPRC), sex was based on self-reported demographic assessment. For the present study, we sought to assess differences between aggregated HC and SZ cohorts. Therefore, we used a general linear model (GLM) approach to control for sex as a covariate in all between-subjects analyses.
Population characteristics	Case (SZ) / Control (HC): 193 / 315 Age (years) mean / standard deviation / median / range (minimum - maximum): 38.48 / 12.94 / 39 / (13 - 68) Sex (female / male): 169 / 339 Race (mixed American (AMR) / European (EUR) / African (AFR) / other): 109 / 304 / 90 / 5
Recruitment	The resting-state fMRI data analyzed in this study was sourced from three case-control psychosis projects: Center for Biomedical Research Excellence (COBRE), Functional Imaging Biomedical Informatics Research Network (FBIRN), and Maryland Psychiatric Research Center (MPRC). COBRE individuals with schizophrenia (SZ) were recruited from the Raymond G. Murphy Veterans Affairs Medical Center and from psychiatric clinics in metropolitan Albuquerque, New Mexico (Aine et al., 2017). SZ individuals received a diagnosis of schizophrenia performed in consensus by two research psychiatrists via the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) using the patient version of the SCID-DSM-IV-TR. SZ subjects were evaluated for comorbidities and for retrospective as well as prospective clinical stability. Additional exclusion criteria were as follows: history of neurological disorder, head trauma with loss of conscious exceeding five minutes, mental retardation, or history of active substance dependence or abuse (except nicotine). Healthy control (HC) individuals from the same geographic location were recruited via Institutional Review Board-approved advertisements and completed the SCID-Non-Patient Edition to exclude individuals with Axis I conditions (Aine et al., 2017). Additional exclusion criteria were as follows: current or past psychiatric disorder with the exception of one lifetime major depressive episode, head trauma with loss of conscious exceeding five minutes, recent history of substance abuse or dependence, occurrence of depression or antidepressant use within the past 6 months, history of antidepressant use exceeding one year, and history of psychotic disorder in any first-degree relative. Individuals did not smoke for at least one hour prior to scanning. FBIRN individuals were recruited across seven different sites within the United States (Damaraju et al., 2014; https://doi.org/10.1016/j.nicl.2014.07.003). SZ individuals were diagnosed with schizophrenia based on the SCID-DSM-IV-TR and

were clinically stable and on antipsychotic medication for at least two months prior to scanning. Exclusion criteria for individuals were as follows: history of major medical illness, MRI contraindications, poor vision with MRI-compatible corrective lenses, IQ under 75, current substance abuse disorder or a history of drug dependence in the last 5 years, extrapyramidal symptoms (for SZ individuals), and current or past history of major neurological or psychiatric illness (SCIS-I/NP) or first-degree relative with Axis I psychotic disorder (for HC).

MPRC individuals with SZ were recruited from outpatient clinics at the Maryland Psychiatric Research Center and mental health clinics in the greater Baltimore area between 2004 and 2016, and HC individuals were recruited via advertisements in the same geographic location (Adhikari et al., 2019). For SZ individuals, a diagnosis of schizophrenia was confirmed via the SCID-DSM-IV. Exclusion criteria for SZ individuals were as follows: major medical or neurological illness, history of head trauma with cognitive sequelae, and diagnosis of intellectual disability. HC exclusion criteria included a past or present diagnosis of DSM-IV Axis I disorder or family history of psychosis in two prior generations.

Case and control participants were compensated for interviews, scan sessions, and assessments conducted during the referenced studies.

Ethics oversight	Subjects provided informed written consent as required and approved by the Institutional Review Boards (IRBs) of the corresponding institutions as follows: COBRE: participants gave written informed consent as required and approved by the IRB of the University of New Mexico (Aine et al., 2017). FBIRN: participants gave written informed consent as required and approved by the IRBs of the University of California Irvine, the University of California Los Angeles, the University of California San Francisco, Duke University, University of North Carolina, University of New Mexico, University of Iowa, and University of Minnesota (Damaraju et al., 2014). MPRC: participants gave written informed consent as required and approved by the IRB of the University of Maryland, Baltimore (Adhikari et al., 2019).

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

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

Life sciences study design

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

Sample size	This study was conducted on resting-state fMRI data collected from three large multi-site datasets. After quality control criteria were used for subject exclusion, the final subject pool included 315 HC and 193 SZ (n = 508).
Data exclusions	The following subject quality control criteria (Iraji et al., 2023; https://doi.org/10.1002/hbm.26472) were used for the current study: 1) completeness of demographic information, 2) availability of T1 structural MRI, 3) availability of genomic information, 4) maximum head rotation less than 3⁰, 5) maximum translation less than 3 mm, 6) mean framewise displacement (FD) less than 0.25, 7) quality registration to an echo-planar imaging template, 8) whole-brain (in addition to the top ten and bottom ten slices) spatial overlap between the subject mask and group mask greater than 80%, and 9) removal of duplicate subjects.
Replication	No replication analysis was conducted for the present study.
Randomization	Participants were organized into case and control groups based on SZ diagnosis (see Recruitment). Between-subjects analyses controlled for covariates including age, sex, site, and motion (mean framewise displacement) using a GLM approach. The independent component analysis (ICA) approach to network estimation included a bootstrapped randomization protocol for each FC metric-specific analysis to ensure component reliability across runs. For the unique ENL ICN validation protocol, ICA was conducted on subsets of 80% of subjects drawn from the total subject pool with five bootstrapped runs per analysis. Statistical randomization (permutation) tests were conducted with 5000 random permutations to assess differences in goodness-of-fit between HC and SZ, differences in estimation reliability between ENL and LIN components, and the robustness of voxel-wise t-tests.
Blinding	No blinding procedure was conducted for the present analysis of the referenced datasets.

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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type
Resting-state fMRI

Design specifications
No task-based fMRI data were included in the present study

Behavioral performance measures
No task-based fMRI data were included in the present study

Acquisition

Imaging type(s)
Functional

Field strength
3T

Sequence & imaging parameters
COBRE data were collected at a single site on a Siemens TIM Trio scanner via an echo-planar imaging sequence (TR = 2000 ms; TE = 29 ms) (Iraji et al., 2022; https://doi.org/10.1162/netn_a_00196). Voxel spacing was 3.75 x 3.75 x 4.5 mm, the slice gap was 1.05 mm, and the field of view (FOV) was 240 x 240 mm. FBIRN data were collected from seven sites (Turner et al., 2013; https://doi.org/10.3389/fnins.2013.00137), with six sites utilizing Siemens TIM Trio scanners and one utilizing a General Electric Discovery MR750 (Iraji et al., 2022). All seven sites used an echo-planar imaging sequence (TR = 2000 ms; TE = 30 ms). Original voxel spacing was 3.4375 x 3.4375 x 4 mm, the slice gap was 1 mm, and the FOV was 220 x 220 mm. MPRC data were collected from three sites via echo-planar imaging sequences (Iraji et al., 2022). One site used a Siemens Allegra scanner (TR = 2000 ms; TE = 27 ms; voxel spacing = 3.44 x 3.44 x 4 mm; FOV = 220 x 220 mm), another used a Siemens TIM Trio scanner (TR = 2210 ms; TE = 30 ms; voxel spacing = 3.44 x 3.44 x 4 mm; FOV = 220 x 220 mm), and the third site used a Siemens TIM Trio scanner (TR = 2000 ms; TE = 30 ms; voxel spacing = 1.72 x 1.72 x 4 mm; FOV = 220 x 220 mm).

Area of acquisition
COBRE field of view (FOV): 240 x 240 mm; FBIRN FOV: 220 x 220 mm (all sites); MPRC FOV: 220 x 220 mm (all sites)

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software
MATLAB, FSL v6.0, SPM 12

Normalization
Nonlinear

Normalization template
EPI template

Noise and artifact removal
Preprocessing was performed primarily within the MATLAB software environment using Statistical Parametric Mapping (SPM12; http://www.fil.ion.ucl.ac.uk/spm/) and the FMRIB Software Library (FSL v6.0; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki). Preprocessing steps included 1) rigid body motion and slice timing correction, 2) nonlinear warping to Montreal Neurological Institute (MNI) 152 coordinate space, 3) spatial resampling to 3 mm isotropic voxel spacing, 4) spatial smoothing with a 6 mm full width at half maximum (FWHM) Gaussian kernel, 5) head motion regression, detrending, despiking, low pass filtering, 6) temporal resampling to TR = 2000 ms, and finally 7) voxel time series Z-scoring to normalize variance.

After the implementation of group-level ICA, components were screened and considered to be artifactual if they exhibited an ICASSO IQ value less than .80, 2) exhibited low visual overlap with gray matter, 3) exhibited peak weight outside of white matter, and 4) exhibited high visual similarity to motion, ventricular, and other known artifacts (Iraji et al., 2023). The subject-level components corresponding to the non-artifactual group-level components were used for further analysis.

Volume censoring
The following subject quality control criteria (Iraji et al., 2023) were used for the current study: 1) completeness of demographic information, 2) availability of T1 structural MRI, 3) availability of genomic information, 4) maximum head rotation less than 3º, 5) maximum translation less than 3 mm, 6) mean framewise displacement (FD) less than 0.25, 7) quality

registration to an echo-planar imaging template, 8) whole-brain (in addition to the top ten and bottom ten slices) spatial overlap between the subject mask and group mask greater than 80%, and 9) removal of duplicate subjects.

Statistical modeling & inference

Model type and settings

Independent component analysis (ICA), univariate parametric and non-parametric analysis, multivariate modeling to balance HC and SZ cohorts for covariates.

Effect(s) tested

No task-based fMRI data were analyzed. The present study tested for the effect of SZ diagnosis on regression of nonlinear whole-brain functional connectivity (NL-wFC) on linear whole-brain functional connectivity (LIN-wFC) goodness-of-fit, effect of explicitly nonlinear (ENL) vs. linear (LIN) connectivity metrics on independent component analysis component estimation reliability (ICASSO IQ), effect of ENL vs. LIN on intrinsic connectivity network (ICN) voxel weight, effect of SZ diagnosis on ENL ICN voxel weight, and effect of SZ diagnosis on LIN ICN voxel weight. We considered age, sex, site, and motion (mean framewise displacement) to be confounding factors and used a GLM approach to account for their effects in all between-subjects statistical analyses.

Specify type of analysis: ☐ Whole brain   ☐ ROI-based   ☒ Both

Anatomical location(s)	We used the Group ICA of fMRI Toolbox (GIFT v4.0; http://trendscenter.org/software/gift) to implement connectivity domain ICA (Iraji et al., 2016; https://doi.org/10.1016/j.neuroimage.2016.04.006) and obtain separate sets of group-level brain networks from linear whole-brain functional connectivity (LIN-wFC) and explicitly nonlinear whole-brain functional connectivity (ENL-wFC) data. The implementation of group-level spatial independent component analysis (gr-sICA) was preceded by an initial subject-level multi-power iteration (Rachakonda et al., 2016; https://doi.org/10.3389/fnins.2016.00017) principal component analysis step to reduce dimensionality and denoise the data (Erhardt et al., 2011; https://doi.org/10.1002/hbm.21170). The 30 principal components that explained the maximum variance of each subject's respective LIN-wFC and ENL-wFC were retained for further analysis. Subject-level principal components from each estimator were concatenated across the component dimension, and a group-level principal component analysis step was applied to further reduce the dimensionality of the data and decrease the computational demands of gr-sICA (Calhoun et al., 2009; https://doi.org/10.1016/j.neuroimage.2008.10.057). The 20 group-level principal components that explained the maximum variance of each estimator-specific data set were used as the input for gr-sICA. We selected a gr-sICA model order of 20 to obtain large-scale functional networks (Iraji et al., 2016; Ray et al., 2013; https://doi.org/10.3389/fnins.2013.00237). To ensure the reliability of our results, ICA was implemented via the Infomax optimization algorithm (Bell & Sejnowski, 1995; https://doi.org/10.1162/neco.1995.7.6.1129) 100 times with both random initialization and bootstrapping, and the most stable run was selected for further analysis. We evaluated the reliability and quality of ENL and LIN components using the ICASSO quality index (IQ), which quantifies component stability across runs (Himberg et al., 2004; https://doi.org/10.1016/j.neuroimage.2004.03.027). To assess the difference in stability between ENL and LIN components, we conducted a two-sided permutation test with 5000 random permutations on the IQ data. Assessing component reliability was a necessary step, as previous work demonstrates that certain components may be inconsistently extracted from the data of interest (Himberg et al., 2004). In the context of fMRI network estimation, ICASSO IQ is often used to differentiate reliable components from components that are unstable and unfit for further analysis (Iraji et al., 2019; https://doi.org/10.1002/hbm.24580). A component was identified as an ICN if and only if 1) it exhibited an ICASSO IQ value exceeding .80, 2) it exhibited high visual overlap with gray matter, 3) it exhibited peak weight within gray matter, and 4) it exhibited low visual similarity to motion, ventricular, and other known artifacts. To find spatially corresponding networks, the spatial correlation value was computed between every pair of extracted LIN and ENL components, and components were matched in a greedy fashion. ICNs matched with a spatial correlation value exceeding .80 were classified as common (Iraji et al., 2023) and were labeled based on their neuroanatomical distributions and the identification of ICNs from previous studies (Iraji et al., 2016). Networks exhibiting maximum spatial correlation less than .40 were classified as unique. We used the Group ICA of fMRI Toolbox (GIFT v4.0) to implement group information-guided ICA (GIG-ICA) (Du & Fan, 2013; https://doi.org/10.1016/j.neuroimage.2012.11.008) and reconstruct subject-specific networks from subject-level principal components using the group-level spatial references. To assess differences in spatial variation between matched networks, we conducted voxel-wise two-sided paired samples t-tests on their Z-scored subject-level estimates. For a given matched network pair, statistical comparisons were masked for voxels exceeding $Z = 1.96$ ($p = .05$) in either group-level map (LIN or ENL), and the False Discovery Rate (FDR) method was used to correct for multiple comparisons ($q < .05$) (Benjamini & Hochberg, 1995; http://www.jstor.org/stable/2346101). The robustness of the voxel-wise t-test procedure was assessed via comparison to the results of voxel-wise two-sided permutation tests with 5000 random permutations for the posterior default mode (pDM) network. The automated anatomical labeling atlas 3 (AAL3) (Rolls et al., 2020; https://doi.org/10.1016/j.neuroimage.2019.116189) was used to localize clusters of significant voxels to anatomically defined brain regions. To assess differences between HC and SZ, we conducted voxel-wise two-sided independent samples t-tests between the estimates of common and unique networks derived from each cohort. We first used a GLM to remove the effect of confounding factors including age, sex, site, and motion (mean framewise displacement) on Z-scored subject-level network estimates. Voxel-wise independent samples two-sided t-tests were then conducted on the residual spatial maps derived from the HC and SZ groups. Statistical comparisons between common networks were masked for voxels exceeding $Z = 1.96$ ($p = .05$) in either of the group-level maps (LIN or ENL), while unique network comparisons were masked for voxels exceeding the same threshold in the unique group-level map. The FDR method (Benjamini & Hochberg, 1995) was used to correct for multiple comparisons ($q < .05$). The robustness of the voxel-wise t-test procedure was assessed via comparison to the results of voxel-wise two-sided permutation tests with 5000 random permutations for the posterior default mode (pDM) network. The AAL3 atlas (Rolls et al., 2020) was used to localize clusters of significant voxels to anatomically defined brain regions. A two-sided McNemar's test was used to assess the overall ENL vs. LIN difference in statistical sensitivity (across all voxels belonging to commonly classified networks), and differences in statistical sensitivity for matched network pairs were investigated separately using either two-sided McNemar's tests or exact binomial tests (for n < 25).
Statistic type for inference (See Eklund et al. 2016)	Voxel-wise
Correction	FDR (Benjamini & Hochberg, 1995; http://www.jstor.org/stable/2346101)

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 (to construct LIN-wFC), distance correlation (to construct NL-wFC), explicitly nonlinear whole-brain functional connectivity (ENL-wFC)

Multivariate modeling and predictive analysis

We used the Group ICA of fMRI Toolbox (GIFT v4.0; http://trendscenter.org/software/gift) to implement connectivity domain ICA (Iraji et al., 2016) and obtain separate sets of group-level brain networks from linear whole-brain functional connectivity (LIN-wFC) and explicitly nonlinear whole-brain functional connectivity (ENL-wFC) data. The implementation of group-level spatial independent component analysis (gr-sICA) was preceded by an initial subject-level multi-power iteration (Rachakonda et al., 2016) principal component analysis step to reduce dimensionality and denoise the data (Erhardt et al., 2011). The 30 principal components that explained the maximum variance of each subject's respective LIN-wFC and ENL-wFC were retained for further analysis. Subject-level principal components from each estimator were concatenated across the component dimension, and a group-level principal component analysis step was applied to further reduce the dimensionality of the data and decrease the computational demands of gr-sICA (Calhoun et al., 2009). The 20 group-level principal components that explained the maximum variance of each estimator-specific data set were used as the input for gr-sICA. We selected a gr-sICA model order of 20 to obtain large-scale functional networks (Iraji et al., 2016; Ray et al., 2013). To ensure the reliability of our results, ICA was implemented via the Infomax optimization algorithm (Bell & Sejnowski, 1995) 100 times with both random initialization and bootstrapping, and the most stable run was selected for further analysis. We evaluated the reliability and quality of ENL and LIN components using the ICASSO quality index (IQ), which quantifies component stability across runs (Himberg et al., 2004). To assess the difference in stability between ENL and LIN components, we conducted a two-sided permutation test with 5000 random permutations on the IQ data. Assessing component reliability was a necessary step, as previous work demonstrates that certain components may be inconsistently extracted from the data of interest (Himberg et al., 2004). In the context of fMRI network estimation, ICASSO IQ is often used to differentiate reliable components from components that are unstable and unfit for further analysis (Iraji et al., 2019). A component was identified as an ICN if and only if 1) it exhibited an ICASSO IQ value exceeding .80, 2) it exhibited high visual overlap with gray matter, 3) it exhibited peak weight within gray matter, and 4) it exhibited low visual similarity to motion, ventricular, and other known artifacts. To find spatially corresponding networks, the spatial correlation value was computed between every pair of extracted LIN and ENL components, and components were matched in a greedy fashion. ICNs matched with a spatial correlation value exceeding .80 were classified as common (Iraji et al., 2023) and were labeled based on their neuroanatomical distributions and the identification of ICNs from previous studies (Iraji et al., 2016). Networks exhibiting maximum spatial correlation less than .40 were classified as unique. We used the Group ICA of fMRI Toolbox (GIFT v4.0) to implement group information-guided ICA (GIG-ICA) (Du & Fan, 2013) and reconstruct subject-specific networks from subject-level principal components using the group-level spatial references.

For the supplementary analysis of balanced HC and SZ cohorts, we used multivariate genetic matching with replacement (Sekhon et al., 2011; https://doi.org/10.18637/jss.v042.i07) to balance cohorts for confounding factors including age, sex, site, and motion (mean framewise displacement).