nature mental health

Perspective

https://doi.org/10.1038/s44220-026-00610-y

Revamping neuroimaging analysis to reveal biomarkers of adolescent mental health

Received: 12 May 2025

Accepted: 17 February 2026

Published online: 02 April 2026

Check for updates

Erica L. Busch[1,2] ✉, Nicholas B. Turk-Browne[1,2] & Arielle Baskin-Sommers[1,2]

Advances in neuroscience research provide an unprecedented opportunity to identify the etiopathogenesis of mental health disorders. Yet it has proven difficult to find reliable associations between neurobiological phenotypes and real-world mental health experiences, particularly among youth. This Perspective addresses two pervasive assumptions inherent to many functional neuroimaging studies that diminish the predictivity of the data. First, studies assume that aligning data across individuals on the basis of the anatomy of the brain is sufficient to align their brain function. Individual brains vary meaningfully in the localization of functions, particularly across development and in clinical populations; neglecting this variability in functional neuroanatomy risks washing out rich and reliable patterns of individual-specific information. Second, studies assume that the underlying signal embedded in brain measurements over space and time can be modeled with simple transformations from high dimensions (that is, voxels) to low or single dimensions (that is, regional averages). However, the latent structure of brain activity and behavior is often complex and nonlinear. To overcome these assumptions, we suggest alternative methodological approaches that have yielded novel insights into the neurobiology of cognition and mental health symptoms in adolescence. Building robust predictive models of psychiatric problems requires methodology that can capture the richness and complexity of the brain and behavior.

Approximately half of the world's population will experience at least one mental health disorder during their lifespan[1]. The majority of such disorders onset during the second decade of life[1]. Adolescents who experience mental health disorders are at greater risk of negative outcomes in adulthood, including chronic mental and physical health challenges, social functioning issues, lower educational attainment, and involvement with the legal system[2–4]. Among people 10–24 years of age, mental health disorders are the leading cause of disability across the world[5]. Accordingly, understanding the neurobiological factors associated with mental health disorders during adolescence has become a major priority of translational research over the past two decades[6–9].

Many aspects of brain functioning have been implicated in the development of mental health disorders, and there have been substantial investments in studying these neurobiological factors across adolescence[10–13]. These studies use mainly functional magnetic resonance imaging (fMRI), which affords non-invasive, whole-brain coverage at a high spatial resolution with minimal risk. fMRI studies have highlighted changes in brain circuitry related to specific psychological states (for example, fear or pain) and cognitive processes (for example, attention or memory) that are commonly altered in mental health disorders[14–19].

These successes in identifying neural signatures of psychological states and processes have inspired confidence that fMRI could be used as a diagnostic tool for mental health disorders. Yet isolating reliable brain-based biomarkers of mental health problems has proven to be far more challenging than anticipated[7,20–25]. Recognized hindrances include the following: (1) small and/or homogeneous samples of participants[17,26,27]; (2) inadequate predictive modeling approaches or parameterization[23,28–31]; and (3) uninformative or unreliable dependent variables (that is, clinical ratings or diagnoses)[32]. Indeed, larger, more

[1]Department of Psychology, Yale University, New Haven, CT, USA. [2]Wu Tsai Institute, Yale University, New Haven, CT, USA. ✉e-mail: erica.busch@yale.edu

diverse samples and more sophisticated analytic approaches have strengthened links between brain measures and individual differences in behavioral measures[23,33,34]. However, some studies with comparable samples and approaches have yielded limited insight into mental health experience, with most studies reporting null or small effects[8,21,35]. Because some individual differences in behavior can be predicted from brain measures in large-scale developmental fMRI studies, the data are clearly not devoid of signal, but appear inadequately posed for clinical applications[36–39].

In this Perspective, we argue that the difficulty of identifying robust relationships between brain function and mental health experiences arises from the choices made by researchers about how to represent their data. We discuss the most common ways fMRI data are prepared and describe data properties that make these steps suboptimal for predicting adolescent mental health. Specifically, we highlight two key assumptions made in many fMRI analyses that limit the field's ability to predict brain–mental health associations. The first assumption is that aligning brain anatomy is sufficient to align brain function. Functional–anatomical correspondence in the brain is highly variable in the general adult population, and even more variable in developmental and clinical populations[40,41]; when this variability remains unaddressed, it limits the informativeness of brain activity for detecting sources of individual differences in behavior across the population[16,40,42]. The second assumption is that linear aggregation methods, which use averaging or other linear combinations to combine fMRI samples across space or time to improve signal quality, are sufficient to capture the complexity of spatiotemporal signals across brain measurements. However, brain activity is better captured with nonlinear models that are robust to noise[29,43–47].

To address these two assumptions, we take guidance from advances in the field of cognitive neuroscience – namely, in the building and application of computational tools to mitigate the noise and variability of fMRI data. By incorporating techniques from machine learning and related fields, cognitive neuroscientists have shown that fMRI signals can be linked robustly with complex behaviors and nuanced cognitive processing, at both individual and group levels[14,16,48]. We anticipate that adapting these computational tools here will afford analogous improvements in the identification of brain-based biomarkers of mental health experiences in the developing brain.

Current approaches in brain–behavior association studies

Collecting fMRI data

fMRI data are collected while participants lie in an MRI scanner and rest or perform tasks that engage specific brain functions. fMRI captures the blood oxygenation level-dependent (BOLD) signal, which is a slow, indirect metabolic measurement of neuronal activity. The spatial unit of measurement for fMRI is a voxel (that is, volumetric pixel) covering approximately 1–3 mm of space in each of 3 dimensions. fMRI measures the brain every 1–2 s, a rate that is substantially slower than neuronal firing. The resulting fMRI volumes are incredibly high dimensional: approximately 100,000 voxels are collected across the entire brain for each of hundreds of samples in a session. Despite the complexity of these signals, fMRI has emerged as the most prominent neuroimaging method for studying brain–behavioral associations for at least three reasons: (1) fMRI generates high-resolution whole-brain images, capturing activity patterns both globally across large-scale brain networks and locally within spatially resolved areas; (2) fMRI is safe and non-invasive, thus accessible to youth and adult participants with a variety of lived experiences; and (3) fMRI is a common tool readily available at research and medical institutes globally[49,50] (Fig. 1).

Establishing population-wide brain activity correspondences

A crucial step in fMRI analysis is establishing a correspondence between brain activity features across participants. This is commonly done by registering the data from all participants to a common anatomical template (for example, Talairach space[51] or Montreal Neurological Institute (MNI) space[52]) on the basis of major anatomical landmarks[53]. However, landmark-based alignment has limitations, particularly in aligning cortical features (for example, sulci, gyri). Another form of anatomical alignment, cortical surface-based alignment, can improve the accuracy of registration by treating the cerebral cortex as a sheet and finding an alignment that matches sulcal and gyral patterns across brains[54]. Compared with landmark-based alignment, surface-alignment approaches more accurately account for morphological and topological properties of the brain, better align anatomy across youth and adults[55,56], and reflect the behaviorally meaningful and predictable ways that cortical anatomy develops[57,58].

Yet even if one could perfectly align brain anatomy across participants, mismatches in brain function would remain. In other words, the same function – whether defined on the basis of information transfer in the brain (that is, functional connectivity) or in terms of the behavioral or cognitive correlates of brain activity – may be performed in slightly different anatomical locations across individuals[59–61]. Researchers commonly mitigate these misalignments in brain function by spatially smoothing the data during preprocessing or by averaging signals across voxels within large, predefined cortical areas. This latter 'region of interest' (ROI) approach blurs signals at the individual level to increase functional overlap across individuals, de-emphasizing fine-grained information specific to an individual in favor of coarse correspondence across individuals. The ROI approach is the basis of most brain–behavior association studies so far[62,63].

Brain measures used for predictive modeling

Traditional ROI analyses extract activity from a brain region selected a priori and relate that averaged signal to behavioral variability across participants, yet cognitive processes commonly implicated in mental health disorders (for example, decision-making or executive control) are related to changes spanning distributed networks of brain regions and their connections. A variety of approaches examine alterations in network-level and whole-brain functional connectivity (that is, the strength of co-activation of different brain areas) associated with mental health experiences[17,20,64,65]. These approaches can be considered multivariate (that is, capturing connections between many pairs of regions), but most also carry assumptions of ROI analyses through the use of parcellation – an atlas that carves up the brain into discrete regions ('parcels'). The activity in each parcel is averaged across voxels and then functional connectivity is commonly calculated as the pairwise correlation between parcels[66–68].

For whole-brain analyses, the parcellation approach offers two advantages. First, as with other ROI approaches, it improves the alignment of functions across brains by allowing for spatial variability in voxel localization. Second, and more specific to parcellation, it reduces the dimensionality of the data from voxels to parcels (3–4 orders of magnitude reduction) in a way that respects the underlying biology that guided the definition of parcels (for example, based on resting-state functional connectivity, histology, or other anatomical features)[65]. Indeed, this dimensionality reduction has made whole-brain analyses more computationally tractable, which is particularly important when working with large sample sizes. Most common parcellations are defined using resting-state functional connectivity from non-clinical, adult brains[65], although developmental and individualized parcellations are gaining traction[69–72].

Predictive models using parcellation-based whole-brain functional connectivity have made progress in understanding individual differences in cognition and behavior among non-clinical youth and adults[38,71,72]. Yet few studies have generated models that successfully predict mental health experiences[25,38,73]. Some researchers have proposed that these models fail because of inadequate training samples[27,28,33], issues with model and parameter selection[27,28,30,74], or low-quality or sparse dependent variables (for example, clinical ratings or diagnoses)[32].

https://doi.org/10.1038/s44220-026-00610-y

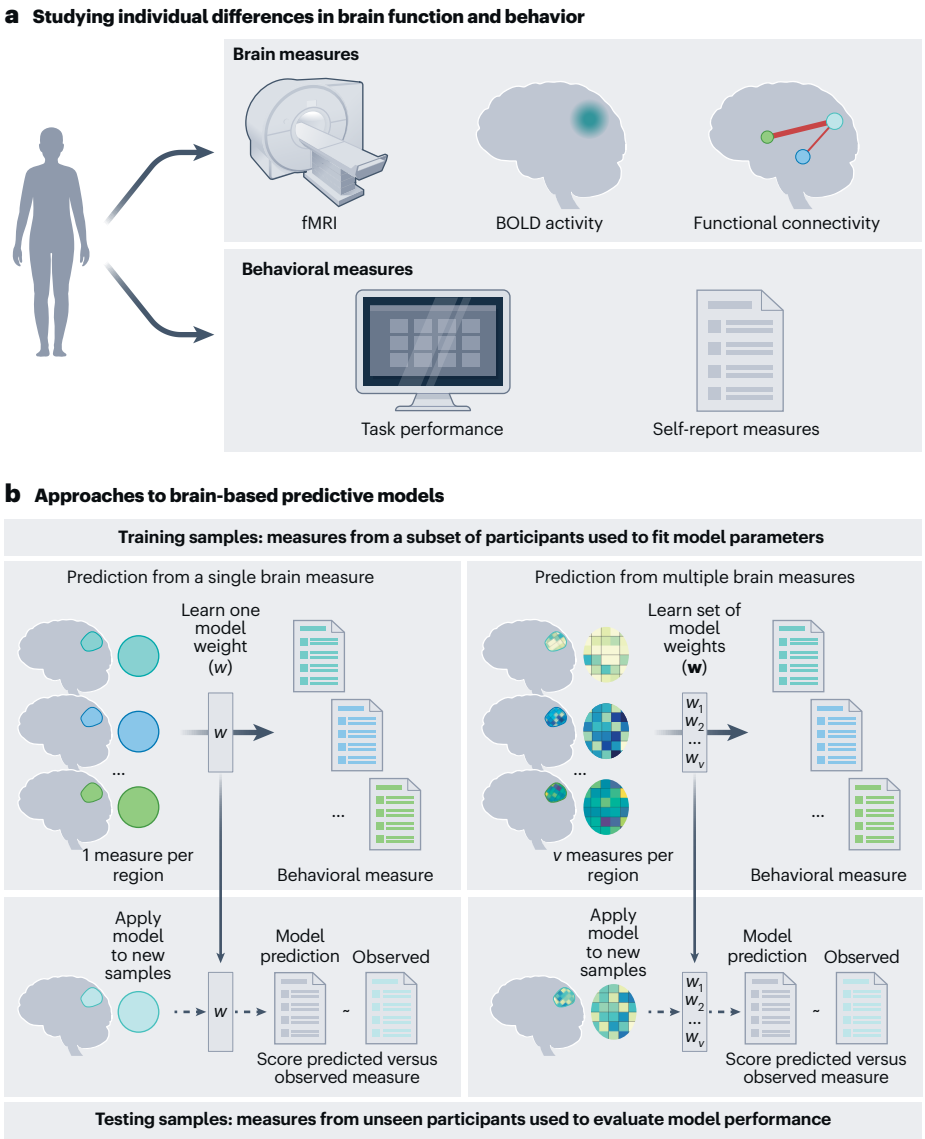

Fig. 1 | Brain–behavior relationships. a, Characterizing relationships between brain and behavioral measures begins with data collection. Brain measures can include activation or functional connectivity and are often collected with fMRI. Common behavioral measures include a battery of cognitive tasks and/or self-report measures meant to characterize a participant's cognitive functioning and/or mental health experiences. **b**, Models that predict behavioral measures from brain measures commonly represent brain activity in one of two ways: a single measurement from averaging across voxels in an ROI (left) or a pattern of measurements from multiple voxels (right). These representations offer two levels of granularity in brain measurements, commonly referred to as univariate and multivariate. Models using univariate information fit a single regression weight per region, and models using multivariate information fit multiple regression weights per region (here, one per voxel). Prediction models should be cross-validated: trained on data from one set of participants and evaluated on data from a separate set of participants.

Rarely has the quality of the independent variable — measures of brain function — been considered the culprit. That is, the way functional brain measures are represented can directly support or hinder prediction of individual differences in mental health experiences. By representing brain activity coarsely with parcels, meaningful variability related to mental health may be washed out. Studies in adults and adolescents have shown that variability in brain activity is most pronounced and reliable at a fine, voxel-scale level of spatial detail[39,75–79] (Fig. 1b). In the following section, we discuss the value of moving to finer spatial scales to reveal reliable individual differences.

Improving alignment of brain function across participants

The first assumption to be addressed is that aligning brain anatomy is sufficient for aligning brain function across individuals. Although functional brain areas and networks are organized similarly across individuals at a coarse level, they vary considerably between individuals at the finer spatial scale of voxels[60,75,76,78]. Individual variability in activity is reliable and predictable within individuals but becomes diluted with group-level aggregation (for example, ROIs, parcellations) and spatial smoothing[39,75,76,78,80]. Furthermore, the information encoded in fine-scale cortical activity affords clearer decoding of psychological states relative to coarse activity[14,79,81]. These reliable signals — commonly obscured by anatomical alignment and averaging — can be retained with functional alignment (Fig. 2).

Functional alignment is a kind of fMRI analysis that maximizes the benefits of both group-level aggregation and fine-scale individual variability. The intuition behind functional alignment stems from the findings that individuals processing the same information (for example, watching the same movie) show synchronized brain activity in many

https://doi.org/10.1038/s44220-026-00610-y

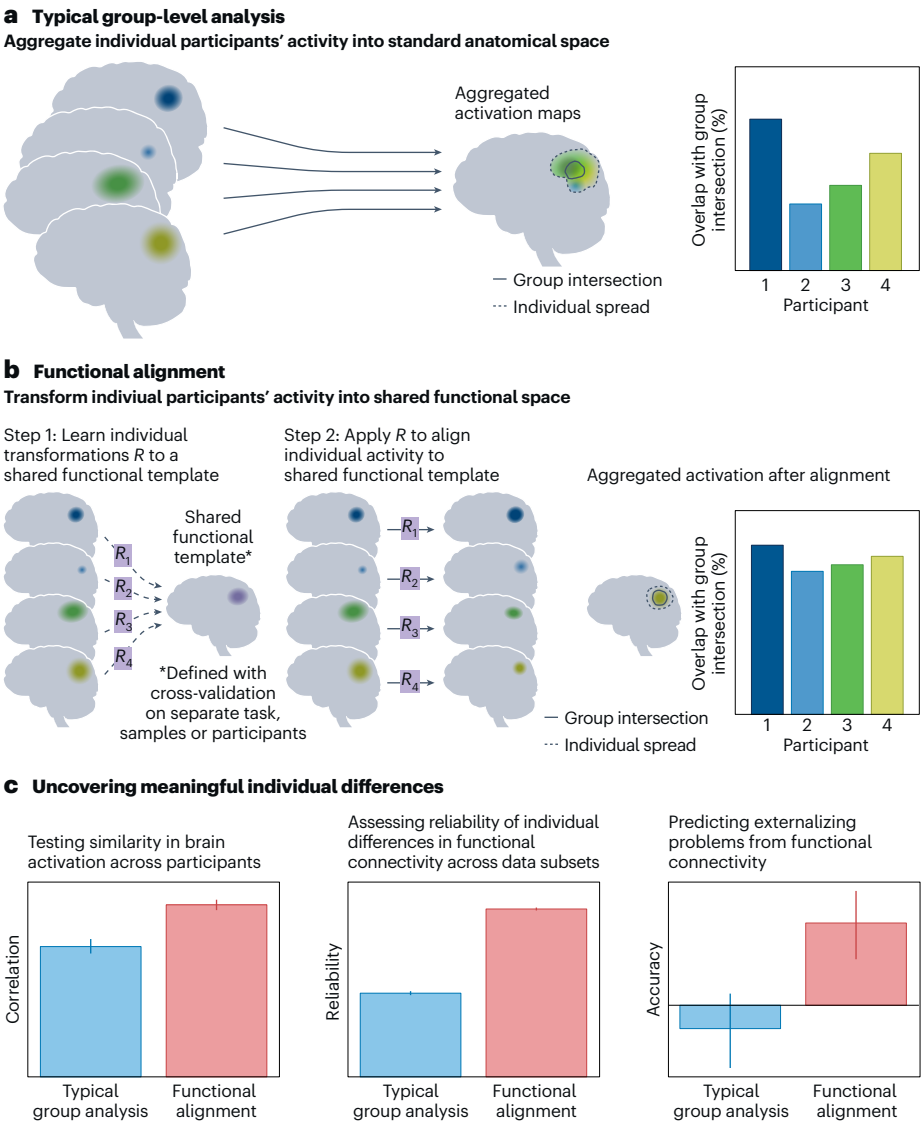

Fig. 2 | Anatomical versus functional alignment. a, The typical approach to group-level fMRI analyses begins with the brain activation maps from all participants aligned to a standard anatomical template. After this anatomical alignment, the activation profiles are unlikely to be fully aligned across participants. Each participant's activation has a different amount and location of overlap with the group intersection, suggesting that modeling only the group intersection misses individual differences. **b**, Functional alignment builds upon the typical group-level analysis after anatomical alignment. The first step is to define a shared functional template, which is determined with cross-validation or on data from separate tasks, samples or participants. This functional template represents brain activity or connectivity patterns shared across participants. There are several algorithms for defining shared functional templates and for deriving transformation matrices to align an individual's brain activity in anatomical space to the shared functional template. This functional alignment leads to greater overlap of participants with the group intersection, while variation across participants in their match to the functional template retains individual differences in the shared space. **c**, Example analyses showing that functional alignment increases: the temporal correlation of cortical responses (left), the reliability of individual differences in functional connectivity across separate runs of resting-state data (middle) and the accuracy of predictions of individual differences in externalizing symptoms (right; unpublished data). Bars represent mean across cross-validation folds; error bars represent the 95% confidence interval of the mean. Panel **c** adapted with permission from: left, ref. 39 under a Creative Commons license CC BY 4.0; middle, ref. 45, Springer Nature Limited; right, E.L.B. et al., unpublished data.

brain regions[82]. Using this evoked synchrony, functional alignment algorithms learn a shared functional 'template' that matches voxels across participants on the basis of the similarity of their activity time courses rather than their anatomical coordinates.

Early functional alignment work showed all participants the same long, time-locked stimulus to determine functional similarity for learning the shared template[60]. This kind of fMRI paradigm is less feasible with adolescents or clinical populations, whereas short (5–10 min) movies or resting-state scans are more accessible[83]. The advantage of using movies (or stories) is that the stimulus can drive functions that may be of interest to align (for example, sensory, affective, cognitive). Although resting-state fMRI does not evoke these common responses, shared patterns of connectivity do emerge at rest and can be used to derive functional templates for alignment[84,85]. Although it may not capture all functions of interest, connectivity-based alignment is more flexible than response-based alignment, as it can be defined from tasks of differing lengths and designs, or from resting-state or sleeping data.

Why would improving the alignment of functions across brains help in the study of individual differences in mental health? It may initially seem counterintuitive that improved alignment would highlight differences in brain activity or connectivity. However, the functional alignment of fine-scale brain measures to a shared template

suppresses undesirable or unreliable sources of variation (for example, in functional–anatomical correspondence, head motion or physiological noise). Thus, the residual inter-individual variation after functional alignment is more reliable than before alignment[39,77,80,86]. These residuals have been used to predict differences in information encoding[77,87], narrative processing[88–91] and measures of cognition[22,80], even among youth[39]. These initial applications suggest that functional alignment could similarly benefit predictive models of mental health experiences[17,22,37].

Implementing functional alignment

Functional alignment is performed on BOLD time-series data typically after basic fMRI preprocessing (for example, using fMRIPrep[92]), including registration to a standard anatomical template (for example, MNI) and regression of nuisance parameters (for example, head motion, physiological signals). There are no specific preprocessing steps 'required' to use functional alignment and optimal parameters will probably come down to the data and analysis goals. Studies with modestly sized adult cohorts have used less intensive nuisance regression protocols[93,94] relative to other studies with larger adult cohorts[80]. Given the complications of developmental neuroimaging and substantial impacts of head motion on functional connectivity estimates within major cortical networks[95], we expect more intensive nuisance regression to be important for developmental applications[39]. In the adult literature, either small (2–4 mm) or no spatial smoothing is performed[93,94]; in fact, increasing smoothing kernel size can diminish the advantages of functional alignment[93]. However, there is insufficient work on functional alignment in adolescents to make definitive recommendations about the ideal amount of spatial smoothing.

The prepared BOLD data are functionally aligned in two steps: (1) the identification of common signals across participants to create a 'shared functional template'; and (2) the calculation of transformation matrices to align each participant's functional signals to this shared functional template (Fig. 2b). The shared functional template and transformation matrices can be computed by providing whole-brain activity patterns or by dividing and conquering with separate templates and transformations for regions of a brain atlas[96] or searchlight analysis[93]. The latter approach of learning and applying transformations to spatially constrained regions, as opposed to the entire brain at once, aids interpretability of functional alignment as the input signals are known to originate from roughly the same anatomical areas across brains. In other words, the remixing of signals happens only at the fine scale and conserves the coarse anatomical structure at the whole-brain level. There are several methods that follow this formula with different computational algorithms and constraints. Below, we outline a few of these approaches, their use cases and strengths, and provide basic implementations (Box 1).

Hyperalignment and high-dimensional models. Consider an experiment in which each participant views the same time-locked, dynamic stimulus (for example, a movie) in the scanner. The resulting fMRI data would comprise the activity of v voxels across t time points (the length of the experiment). During hyperalignment, the pattern of brain activity across voxels at each time point can be formulated as a vector in a v-dimensional space. Each axis v_i of that space is a voxel and the vector's coordinate on v_i is the BOLD amplitude of the corresponding voxel at a given time point. The next time point is also a vector, and so on, resulting in a trajectory of t vectors (one for each time point) through the v-dimensional space over time. This trajectory represents the coordinated activity of individual voxels across the course of the experiment.

This vectorization is different than standard fMRI analysis approaches that treat brain activity as one-dimensional (that is, a single voxel or an ROI average) or three-dimensional (that is, the brain's x, y, z anatomical coordinates). Instead, each participant has their own v-dimensional information space, and their trajectory through this space defines the geometry of their brain states while viewing the stimulus. Because all participants viewed the same stimulus, brain regions related to processing that stimulus may exhibit a similar brain-state geometry across participants. In turn, this enables the identification of a shared v-dimensional space that results from the transformation of each participant's space. In hyperalignment, the transformations are identified with singular value decomposition, which minimizes the Frobenius norm between a source and target matrix (for example, a participant's data matrix and the shared functional template). The resulting transformations minimize the distance of activity vectors for corresponding time points across participants while maintaining the distance between vectors over time within participant, maximizing the similarity of participants' trajectories[40,60].

Hyperalignment achieves this objective using generalized Procrustes analysis to define a group-level shared functional template and individual transformations to map participants to this template. The Procrustean transformation is an orthogonal transformation (that is, allows rotations and reflections) that minimizes the Euclidean distance between two data matrices with the same number of samples t and features v (for example, fMRI datasets with the same number of time points and voxels). These transformations are derived by optimizing for maximal similarity in the representational geometry, such as correlations or covariance structures, between the neural activity patterns of different individuals. Generalized Procrustes analysis finds these transformations iteratively at the group level, by aligning each new participant's data matrix to the average of the previously aligned matrices, which serves as the template. After this iterative identification of the template, a final set of Procrustes transformations are derived to align data from each participant's data to this final template[40].

Hyperalignment can be performed with any brain measure with a geometric representational structure. As noted earlier, although traditionally used to align stimulus-evoked voxel responses over time (that is, when viewing the same movie or story)[60,77,93], hyperalignment has also been used to align functional connectivity[84]. The vectors in the v-dimensional space represent connectivity strength with each seed voxel or ROI. As before, each axis v_i of this space is a voxel, but the vector's coordinate on v_i is the Pearson correlation coefficient (or other similarity/distance metric) of voxel i's BOLD time series with that of the seed voxel or region. In this case, instead of aligning the geometry of brain-state trajectories, hyperalignment identifies a shared v-dimensional space that aligns the geometry of functional connectivity. In practice, this often results in improvements for downstream analyses comparable to aligning based on synchronized stimuli, because connectivity patterns are rich, consistent within an individual and informative at the group level[84,94].

After hyperalignment identifies a shared functional template, data from new participants can be mapped to that template, enabling robust cross-validation and assessment of generalization[39,97,98]. Alignment to the shared template increases the reliability and behavioral relevance of differences between an individual's brain and the group, as shown in resting-state and tasked-based functional connectivity in adults[80] and children[39]. In addition, the weights of the transformation matrices from a participant's brain to the shared template can capture mental health-related individual differences (for example, transdiagnostic biomarkers of psychosis[86]). Hyperalignment is an anatomically interpretable method, as the shared template and transformation matrices retain the input (for example, voxel) dimensions, so the hyperaligned data can be projected back onto the brain to evaluate where functional differences originate.

Shared response model and low-dimensional solutions. The Procrustes-based hyperalignment approach generally assumes a high-dimensional shared functional template, where k, the dimensionality of the resulting model, matches the number of inputs (that

BOX 1

Implementation resources

We aim to make the approaches introduced in this Perspective easy to apply on a range of neuroimaging data types using open-source software. Here we provide a practical 'starter guide' for applying these methods. We also provide sample fMRI datasets and simulated data to replicate analyses from prior papers. These tools are hosted on GitHub (github.com/ericabusch/RepBiomrkr) with instructions for running the sample code on Google Colab or locally as Jupyter notebooks in Python.

Functional alignment

The functional alignment notebook covers hyperalignment[60] and the SRM[87]. We include a hyperalignment implementation in the provided hyperalignment.py Python script, which we demonstrate converges with the implementation from the nltools package (github.com/cosanlab/nltools). We use the implementation of SRM from the publicly available BrainIAK package (brainiak.org). Hyperalignment and SRM are demonstrated for the following use cases:

(1) Simulated data. We provide a package for simulating multivariate datasets with a known correlational structure, both within and across datasets. This allows us to mimic properties of fMRI datasets with differing levels of latent signal structure and noise, and to test the effects of different alignment procedures on improving between-dataset correspondence.

(2) Movie-viewing data. We provide preprocessed fMRI data from an open-source dataset (arks.princeton.edu/ark:/88435/dsp01nz8062179) in a visual ROI as an example. We use intersubject correlation[82] as a metric to demonstrate how functional alignment improves the correspondence of time-locked signals across participants. We also demonstrate how functional alignment improves between-subject classification, and how hyperalignment and SRM transformations and shared functional templates can be cross-validated and applied to out-of-sample data.

(3) Resting-state functional connectivity. We provide code that accesses an open-source resting-state fMRI dataset, selects a seed ROI, and computes the functional connectivity between the seed region and target regions across the brain (http://nilearn.github.io/stable/modules/generated/nilearn.datasets.load_nki.html). We show how functionally aligning these signals improves the reliability of individual differences in the data; that is, after functional alignment, functional connectomes are more distinct across participants but more reliable within participant across split halves of data.

Manifold learning

The manifold learning notebook covers many approaches to dimensionality reduction, including PCA, LLE (locally linear embedding), Isomap (isometric mapping algorithm), *t*-SNE (*t*-distributed stochastic neighbor embedding), PHATE (potential of heat diffusion for affinity-based transition embedding) and E-PHATE (exogenous PHATE). We demonstrate these algorithms with the following use cases:

(1) Iris plants dataset. This is a simple, classic classification dataset available from sklearn: scikit-learn.org/stable/datasets/toy_dataset.html#iris-dataset. The dataset contains three classes (each a type of iris plant) with 50 samples each, for a total of 150 samples with 4 features each. One class is linearly separable from the other two; the latter two are not linearly separable from each other and you can visually see this in the embedding visualizations. This dataset is also available via the UC Irvine Machine Learning Repository (archive.ics.uci.edu/dataset/53/iris).

(2) Handwritten digits dataset. This dataset is a bit more complicated, containing images of handwritten digits (10 classes, where each class refers to a digit). The data are 64-dimensional (8×8-pixel images, vectorized), so it is higher-dimensional and noisier than the Iris dataset. In the embedding spaces, you can see how linear dimensionality reduction (PCA) does not separate the classes, whereas two nonlinear methods do. This dataset is also accessible via sklearn: scikit-learn.org/stable/datasets/toy_dataset.html#optical-recognition-of-handwritten-digits-dataset or from the UC Irvine Machine Learning Repository (archive.ics.uci.edu/dataset/80/optical+recognition+of+handwritten+digits).

(3) Micro-mass dataset. This dataset is the most like what we would see with fMRI: high-dimensional biological data with far more features than samples. The features are mass-spectrometry metrics for 10 classes of microorganisms, with 36 examples for each class, for a total of 360 samples with 1,301 features per sample. Visualizations of this dataset show a clear instance in which nonlinear embedding methods outperform linear methods in distinguishing the classes. This dataset is freely available on OpenML (openml.org/search?type=data&id=1514) and accessible via the fetch_openml function from sklearn (scikit-learn.org/stable/modules/generated/sklearn.datasets.fetch_openml). This dataset is also available from the UC Irvine Machine Learning Repository (archive.ics.uci.edu/dataset/253/micromass).

(4) Simulated fMRI dataset. We created a simulated dataset of 100-dimensional brain measures for 400 participants, a 6-dimensional simulated exogenous data matrix for the 400 participants, and a single variable of interest to predict. These data were generated to highlight the effects from a previously published paper[46], where the 100 dimensions are simulated voxels, the 6 dimensions are simulated environmental factors, and the variable to be predicted is a simulated behavioral measure. Note that these data were generated to magnify the effect of exogenous features in the prediction problem for educational purposes and may show larger effects than with real-world data. We show how embedding these data with the different approaches reflects individual differences in participant scores. We also provide sample code for running cross-validated prediction analyses on embedded data. These data are available via our GitHub repository: github.com/ericabusch/RepBiomrkr/tree/main/sample_data.

is, voxels) v. The number of dimensions k of the shared template in hyperalignment can be reduced from v to capture shared information in lower dimensions, reduce noise and prevent overfitting[60]. This is often performed with principal component analysis (PCA) over the shared template, after fitting the voxel dimensions. Notably, using PCA does not outperform the high-dimensional solution, but does improve computational tractability[60,93].

The shared response model (SRM) approach formulates its shared template as an inherently low-dimensional signal space. That is, it pre-specifies a lower dimensionality ($k < v$) for the shared template, where the dimensionality k can be tuned as a hyperparameter[87]. SRM uses a probabilistic latent-factor model to decompose each participant's BOLD activity into two matrices: a low-dimensional shared response matrix $S \in \mathbb{R}^{k \times t}$ that is common across participants (that is,

the shared functional template) and a participant-specific, orthonormal basis matrix $W_i \in \mathbb{R}^{v \times k}$ (that is, the transformation matrix from v voxels onto k features for participant i). These two matrices are fit jointly over a number of iterations. First, the transformation matrices W_i and shared response S are randomly initialized, and they are jointly optimized using an expectation-maximization scheme. After being fit, a new participant j who completed the same task can be mapped to S to derive their own transformation matrices W_j. Once the transformation weights have been determined, they are re-used to align different samples (for example, separate time points or tasks) from the same participants[87].

SRM has been used to align features of brain activity driven by a common stimulus in a variety of naturalistic paradigms (for example, movies or narratives). In this aligned space, individual differences related to stimulus processing also become more pronounced. For example, SRM has helped reveal features of stimulus-evoked brain activity related to cognitive development[99], affective processing[100], event memory[90], semantic representations[91] and psychological states such as paranoia[88]. Within psychiatry, a few studies have applied SRM to distinguish brain pediatric anxiety[101], and craving and recovery from substance use disorders[102].

Although alignment with SRM can highlight how individuals differ from the group model, it may be suboptimal for instances of substantial, phenotypically driven variability (for example, comparing infants with adults) because of its objective of optimizing a shared response and warping participants to fit to that template[16]. Robust SRM addresses this limitation by adding another factorization to the separate shared and individual components of brain activity. In robust SRM, the individual components are sparse and therefore assume infrequent individual deviations from the group[103]. Like hyperalignment, the family of SRM approaches — which traditionally leverage data collected with time-locked stimuli — can be generalized to data represented as functional connectivity[85].

Owing to its low-dimensional template formulation, SRM can be used more flexibly than hyperalignment, such as when the signals to be aligned have different dimensionality because of variation in brain volume or the use of personalized networks or parcellations[71]. In this case, personalized networks may account for individual differences in the localization of brain activity covariance, whereas SRM could account for individual differences in cortical responses within those networks. SRM also relaxes the assumption that brains should start off aligned anatomically, which makes SRM useful for functional alignment across fundamentally different brains, such as stages of development, species or neurological diseases.

Overall, the early applications of these methods suggest that functional alignment will be useful for studying individual differences in mental health experience in three ways. First, functional alignment can be used as a preprocessing step to address functional–anatomical mismatches at a population level[39]. Second, the shared templates can be analyzed across different demographics or diagnostic criteria[99,101,102]. Finally, the transformation weights between individuals and a shared template can be assessed to gauge topographic shifts in the anatomical localization of brain functions that relate to behavioral or experiential variation[86].

Uncovering the low-dimensional structure of functional signals

Modeling latent signals from high-dimensional data

The second assumption of many fMRI analyses is that linear aggregation methods well approximate brain signal structure. fMRI produces enormously high-dimensional data: brain activity measured in approximately 100,000 voxels at each time point, with often hundreds of time points collected per scan. One of the main advantages of fMRI is that, compared with other non-invasive neuroimaging methods (for example, electroencephalography), fMRI voxels have higher spatial resolution (typically 1.5–3 mm isotropic), allowing for fine-grained analyses of brain activity in vivo. Yet even with the smallest voxel size possible, each voxel captures the activity of thousands of neurons. Furthermore, the BOLD signal measured at each voxel is orders of magnitude slower than the spiking activity and electrical potentials of neurons underlying a vascular response. The substantial spatial and temporal autocorrelation of neuronal signals measured by fMRI makes individual measurements across space or time redundant and explainable in far fewer, simpler dimensions than initially measured during data acquisition.

Redundancy is a common property of neural population activity[104]. Consider recording from a population of 100 neurons that contribute to behavior y. Given synaptic connections among those neurons, the firing of each neuron is not independent of the others; their firing rates would co-vary and overlap in ways such that y can be predicted by k signals, where $k << 100$ neurons. These k signals are latent in that they are not directly measurable (that is, they are not simply the activity of k specific neurons), but can be modeled mathematically by considering the patterns by which the neurons fire together (that is, their covariance). Using fMRI to measure neural activity adds an additional layer of complexity because fMRI as a technique introduces spatial and temporal artifacts, many of which are nonlinear. Thus, simplifying or reducing the dimensionality of fMRI data to model its latent signals can help sift through the various sources of measurement noise and neuronal co-modulation to access a clearer representation of a brain area's overall activity.

Linear dimensionality reduction

The most common approach to simplifying high-dimensional neuroimaging data is with ROIs or parcellations, as discussed earlier. They reduce complexity by downsampling the data in space, but they assume that each region has homogeneous signals and that the boundaries of an ROI or parcel reflect meaningful functional boundaries. Within a single region, voxels may exhibit heterogeneous patterns of activity that reflect distinct functional subnetworks or interactions with other regions[65,76,105]. Averaging across voxels may obscure individual differences in this fine-grained structure of neural activity that may be relevant for understanding mental health experiences[50]. This limitation highlights the need for dimensionality reduction methods that can capture the richness of voxel-level interactions without collapsing them into a single summary statistic (Fig. 3).

Unsupervised dimensionality reduction techniques, such as PCA and independent component analysis (ICA), offer nuanced ways to extract features from high-dimensional data. PCA identifies orthogonal components that explain the maximum variance in the data, making it useful for identifying dominant activity patterns. ICA decomposes the data into statistically independent signals, which can reveal distinct functional networks or sources of signal[64,106]. These methods provide simplified multivariate representations of brain measures and have yielded insight into cognition from fMRI, with limited successful extension to clinical variables[25,31,107–109].

Supervised dimensionality reduction techniques, such as canonical correlation analysis or partial least squares regression, have been used to reduce dimensionality of brain measures to maximize association with mental health variables[110,111]. Because the dimensionality reduction is trained with supervision (that is, optimizing for the prediction goal), the resulting embedding may not reflect the intrinsic latent structure of the brain measures that could arise naturally through unsupervised methods. Nevertheless, these supervised methods (canonical correlation analysis and partial least squares regression) are similar to the unsupervised methods above (PCA and ICA) in that they assume linear interactions among the brain measures (that is, that the activity of different voxels is linearly related).

Nonlinear dimensionality reduction

Recent work has shown that more complex behaviors are best captured by nonlinear interactions between brain measurements. This may be

https://doi.org/10.1038/s44220-026-00610-y

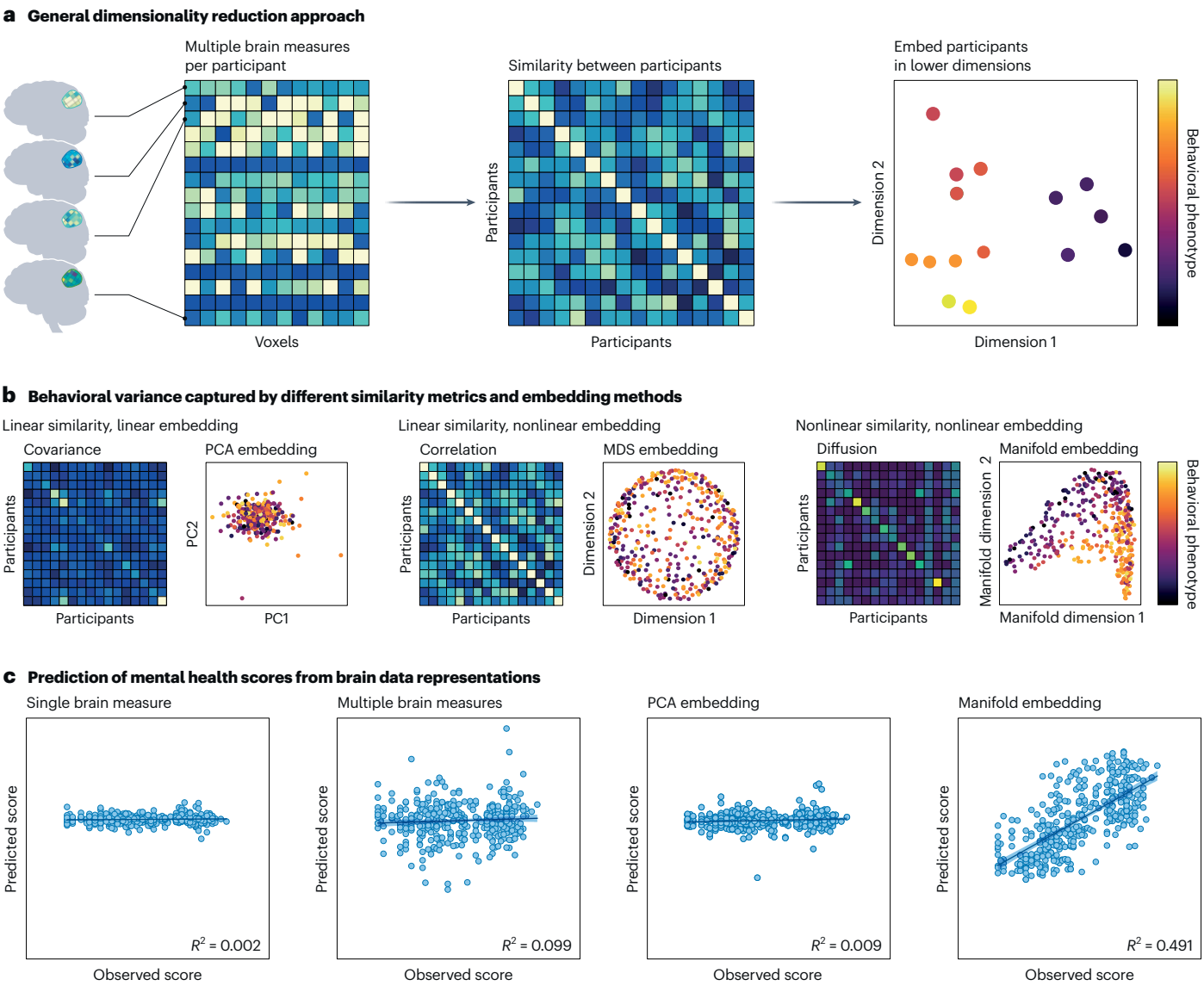

Fig. 3 | Simulations of dimensionality reduction. The data used in this figure were simulated to clearly illustrate known effects of the relationship between biological and environmental variables in predicting mental health outcomes[46] (Box 1). **a**, Dimensionality reduction pipelines take multiple brain activity measures per participant (Fig. 1b) and formulate them into a matrix with participants as samples and brain measures (here, voxels) as features. Pairwise similarities between participants' brain measures are summarized in an affinity matrix. These pairwise affinities are used to project (or 'embed') brain measures into a lower-dimensional representation (that is, a participants-by-dimensions matrix; here, with two dimensions for visualization). A meaningful embedding would ideally show participants clustered according to a label of interest. Crucially, these labels are never used in defining the embedding; the clustering emerges as a facet of modeling the data well. **b**, Left, PCA defines affinity as the covariance between brain measures then projects the samples onto the first k eigenvectors of this covariance matrix (here, $k = 2$ for visualization). Middle: multi-dimensional scaling (MDS) defines similarity as correlation distance or Euclidean distance and embeds the data to minimize the overall difference of pairwise distances between points in the embedding space and in the affinity matrix. Right, manifold learning (here, with E-PHATE) uses nonlinear computations (for example, potential heat diffusion) to summarize the affinity between participants and embeds the data with a nonlinear projection. **c**, Prediction of mental health scores from each data representation. All representations begin with the same brain measures (shown in **a** for the first 16 participants). Ridge regression models were trained with cross-validation (Fig. 1b) to take single brain measures (first panel; average across voxels), multiple measures (second panel; voxel resolution as in **a**), dimensionality reduced measures with PCA (third panel) or E-PHATE (fourth panel) and learn a relationship with mental health scores. In held-out data, observed scores were compared with the model's predicted scores from each data representation. Manifold embeddings provided the best representation for trained models to predict of mental health scores.

a reflection of nonlinear operations performed during neuronal processing and information transfer. Although linear interactions are often easier to interpret, nonlinear interactions are more biologically valid, and so capturing these nonlinearities allows the shape of the data to be more faithfully represented[47,112,113]. A family of nonlinear approaches called 'manifold learning' — which (broadly) attempt to generalize linear dimensionality reduction frameworks to be sensitive to nonlinear structure in data — may therefore be better equipped for the nature and noise of brain activity measures.

Manifold learning has become increasingly popular over the past decade and has been used to predict disease outcomes from high-throughput biomedical data (for example, single-cell RNA sequencing and flow cytometry panels)[114–116]. Like unsupervised linear methods (for example, PCA and ICA), manifold learning approaches

have the goal of discovering a low-dimensional, denoised projection of the data without the use of pre-determined labels. Common manifold learning algorithms include diffusion maps[117], UMAP (universal manifold approximation and projection)[115], *t*-SNE[118], LLE[119] and PHATE[120]. These manifold learning algorithms uncover additional structure that enables the prediction of nuanced disease information, such as individual patient outcomes from their flow cytometry patterns[116].

The recent success of manifold learning in characterizing complex biomedical data and predicting individual differences in disease experience and outcome suggests its utility in the realm of mental health. We envision applying manifold learning as a final step of data preparation before analysis. We have found that the PHATE method is robust to many fMRI preprocessing and denoising techniques, including detrending, filtering and nuisance regression, as well as spatial smoothing, although this robustness would need to be assessed further in a developmental cohort. Recent work has applied manifold learning to fMRI data in various ways: to uncover individual- and region-specific dynamics related to cognitive processing[45], to model group-level, whole-brain functional connectivity patterns across rest and tasks with varying cognitive loads[43,121,122], and to model individual differences in brain activation related to cognitive and emotional processing during development[46]. These findings show the versatility of manifold learning (relative to linear dimensionality reduction) when applied to time-series fMRI data, whole brain or regional functional connectivity, task activation weights, and data at the group or individual level[43,44,46,122].

A particular strength of manifold learning is its ability to incorporate multiple data views (that is, measurement types). Multi-view manifold learning allows for nonlinear representations of several unrelated feature sets from matched samples to be weighted and combined into a single representation. For example, a single-view approach would consider how a gene expression pattern X results in disease phenotype y, but a multi-view approach could consider how gene expression X_1 interacts with patterns of protein concentrations X_2 to predict y. In other words, multi-view manifolds provide more comprehensive accounts of the biological processes that inform outcomes because they consider the interaction of information from diverse sources[123].

We initially developed a multi-view manifold learning algorithm to characterize the multiple levels of signal properties endogenous to fMRI data that were related to cognitive processes (for example, perception, narrative comprehension). This temporal PHATE (T-PHATE) approach provides two views of fMRI time series: one characterizing the interactions among voxels at each time point, and one modeling the temporal dynamics in each voxel. Combining these two views results in a single representation of the brain activity that incorporates interactions between both signal properties over space and time, which proved essential for unveiling how the brain moves through different states during an experiment[45]. T-PHATE embeddings provide rich, cleaned representations of participant-specific and behaviorally relevant brain activity dynamics in lower dimensions. Furthermore, experiment-driven behavioral and cognitive processes were more clearly reflected in the T-PHATE embeddings of brain activity than in embeddings based on a variety of linear and nonlinear dimensionality reduction methods or ways of including temporal dynamics in a single-view manifold[45].

We then generalized our multi-view manifold learning approach to enable the modeling of individual differences in behavior and mental health as combination of endogenous measures (for example, biological features) and exogenous measures (for example, environmental features). Many facets of an adolescent's life inform current and future mental health, and, like neurobiology, environments are high-dimensional, complex systems that have nonlinear relations with mental health experiences[124–127]. We combined measures of adolescents' brain activity and measures of their environment (for example, family and neighborhood adversity) into a single E-PHATE manifold. This allowed us to test how family and neighborhood features interact with brain function and predict mental health experience. E-PHATE embeddings that combined brain activation with environmental information (separately for several subcortical regions and cortical networks) strongly predicted overall mental health problems and externalizing and internalizing symptoms cross-sectionally[46].

These E-PHATE embeddings yielded higher accuracy than other approaches — including spatial averaging, dimensionality reduction with linear or other nonlinear methods, or retaining high-dimensional measures (voxel) — used in the same brain areas and networks. Moreover, E-PHATE embeddings yielded selective longitudinal insight into adolescent mental health experiences: future externalizing symptoms were moderately predicted from embeddings of ventral attention network activation, and future internalizing symptoms were moderately predicted from embeddings of both frontoparietal and ventral attention network activation[46]. The effect size of longitudinal prediction was weaker than within-time-point prediction, but was statistically significant and interpreted as average effect size for a large developmental sample[128,129]. Furthermore, longitudinal prediction revealed specific and consistent cortical networks that were implicated across time points[46].

Together, these findings show that individual differences in brain activity related to cognition, behavior and mental health experiences can be modeled as variability along a low-dimensional brain activity manifold. Nonlinear manifolds are particularly robust to high-dimensional data with high noise, motion and artifact, making them ideal for these pervasive features of developmental neuroimaging data. E-PHATE, a general purpose multi-view manifold learning algorithm, allows us to combine multiple data sources from the same participants to gain a more holistic picture of factors that inform mental health risk at present and in the future[46].

Important considerations with manifold learning include the following: (1) the difficulty of cross-validating nonlinear embeddings or extending latent spaces to samples not seen during training; and (2) the challenge of inverting samples from their embeddings back to the input space (that is, specific brain regions or voxels). We recommend applying manifold learning within spatially constrained regions or networks for questions pertaining to the contribution of specific areas to a behavioral phenotype. Manifold learning can be applied to whole-brain data, but interpreting the contributions of individual brain areas within the embedding is a challenge due to the nonlinear transformation from the brain to the embedding. Improving interpretability and extensibility are active areas of investigation in manifold learning and representation learning more generally[130–132].

Discussion

A major goal of translational neuroimaging research is to establish robust brain-based predictors of adolescent mental health[6,7,17]. The urgency of addressing mental health issues in adolescence has prompted substantial initiatives for neuroimaging data collection[12,48]. These initiatives have deepened our understanding of the brain measures underlying some of the cognitive-affective mechanisms associated with various mental health conditions. Yet the clinical utility (that is, ability to detect or predict variability related to symptoms or diagnoses) of such brain measures has remained limited[7,21,35]. Previous work has focused on this gap being a result of low data quantity and diversity (that is, number of participants and their sociodemographic distribution), predictive modeling techniques, or the reliability of phenotypical measures[17,23,48]. Here we instead focus on the methodological constraints inherent to preparing brain measures for individual differences-style analyses.

We outlined two recent computational methods from cognitive neuroscience — functional alignment and manifold learning — that can enhance the signal quality, between-participant alignment and behavioral relevance of brain measures collected with fMRI in a data-driven fashion. Functional alignment addresses the problem of

mismatches between participants in what a brain signal encodes and where it is encoded anatomically. This provides access to true sources of individual differences in brain function, disentangled from anatomical variability or noise sources of less interest[16,40]. Manifold learning relaxes the assumption that the signal structure of brain activity will be well captured in linear dimensions. Embedding brain measures into lower, nonlinear dimensions allows the natural shape of the data to emerge[46,116].

Initial applications of these methods have yielded refined representations with substantially clearer associations between brain measures and behaviors among adolescents[39,46]. In the future, we anticipate that these more precise and informative representations can be used to identify reliable brain-based biomarkers of different symptoms and levels of mental health disorders.

Furthermore, manifold learning is a method that can answer the call by psychologists to confront the role of both biological and environmental variables in the onset and maintenance of mental health disorders[124–127,133]. Manifold learning handles the noise inherent in high-dimensional brain measures while incorporating the nonlinear effects of exogenous risk factors on the brain, providing a holistic representation of the individual variability related to mental health. Refinement of this method to incorporate estimates of nonlinear change over time in both brain and environmental measures is an exciting opportunity to capture the biosocial transactions that inform mental health experiences across development[127].

We hope that further development and proliferation of these tools will shed light on specific levels and combinations of factors associated with various mental health experiences. The approaches covered in this Perspective and future approaches will help researchers take full advantage of large neuroimaging datasets and realize their intended clinical utility by offering refined representations of adolescent neurobiology and environments. Clinical psychological science is a powerful tool for identifying markers of, and intervening on, mental health disorders. However, psychological scientists need not act alone. Instead, progress will benefit from cooperation and collaboration among different types of scientist, clinical and computational, to derive accurate biomarkers of mental health disorder risk and potential targets for intervention.

References

1. McGrath, J. J. et al. Age of onset and cumulative risk of mental disorders: a cross-national analysis of population surveys from 29 countries. *Lancet Psychiatry* **10**, 668–681 (2023).
2. Johnson, D., Dupuis, G., Piche, J., Clayborne, Z. & Colman, I. Adult mental health outcomes of adolescent depression: a systematic review. *Depress. Anxiety* **35**, 700–716 (2018).
3. Clayborne, Z. M., Varin, M. & Colman, I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. *J. Am. Acad. Child Adolesc. Psychiatry* **58**, 72–79 (2019).
4. Copeland, W. E., Alaie, I., Jonsson, U. & Shanahan, L. Associations of childhood and adolescent depression with adult psychiatric and functional outcomes. *J. Am. Acad. Child Adolesc. Psychiatry* **60**, 604–611 (2021).
5. Gore, F. M. et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. *Lancet* **377**, 2093–2102 (2011).
6. Insel, T. R. & Quirion, R. Psychiatry as a clinical neuroscience discipline. *JAMA* **294**, 2221–2224 (2005).
7. Gabrieli, J. D. E., Ghosh, S. S. & Whitfield-Gabrieli, S. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron* **85**, 11–26 (2015).
8. Etkin, A. A reckoning and research agenda for neuroimaging in psychiatry. *Am. J. Psychiatry* **176**, 507–511 (2019).
9. Uddin, L. Q., Castellanos, F. X. & Menon, V. Resting state functional brain connectivity in child and adolescent psychiatry: where are we now? *Neuropsychopharmacology* **50**, 196–200 (2024).
10. Satterthwaite, T. D. et al. Neuroimaging of the Philadelphia Neurodevelopmental Cohort. *NeuroImage* **86**, 544–553 (2014).
11. Di Martino, A. et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* **19**, 659–667 (2014).
12. Casey, B. J. et al. The Adolescent Brain Cognitive Development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
13. Horien, C. et al. A hitchhiker's guide to working with large, open-source neuroimaging datasets. *Nat. Hum. Behav.* **5**, 185–193 (2021).
14. Haxby, J. V., Connolly, A. C. & Guntupalli, J. S. Decoding neural representational spaces using multivariate pattern analysis. *Annu. Rev. Neurosci.* **37**, 435–456 (2014).
15. Rosenberg, M. D. et al. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat. Neurosci.* **19**, 165–171 (2016).
16. Cohen, J. D. et al. Computational approaches to fMRI analysis. *Nat. Neurosci.* **20**, 304–313 (2017).
17. Woo, C.-W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nat. Neurosci.* **20**, 365–377 (2017).
18. Taschereau-Dumouchel, V., Kawato, M. & Lau, H. Multivoxel pattern analysis reveals dissociations between subjective fear and its physiological correlates. *Mol. Psychiatry* **25**, 2342–2354 (2020).
19. Zhou, F. et al. A distributed fMRI-based signature for the subjective experience of fear. *Nat. Commun.* **12**, 6643 (2021).
20. Rosen, B. R. & Savoy, R. L. fMRI at 20: has it changed the world? *NeuroImage* **62**, 1316–1324 (2012).
21. Finn, E. S. & Todd Constable, R. Individual variation in functional brain connectivity: implications for personalized approaches to psychiatric disease. *Dialogues Clin. Neurosci.* **18**, 277–287 (2016).
22. Finn, E. S. & Rosenberg, M. D. Beyond fingerprinting: choosing predictive connectomes over reliable connectomes. *NeuroImage* **239**, 118254 (2021).
23. Greene, A. S. et al. Brain–phenotype models fail for individuals who defy sample stereotypes. *Nature* **609**, 109–118 (2022).
24. Blair, R. J. R., Mathur, A., Haines, N. & Bajaj, S. Future directions for cognitive neuroscience in psychiatry: recommendations for biomarker design based on recent test re-test reliability work. *Curr. Opin. Behav. Sci.* **44**, 101102 (2022).
25. Jirsaraie, R. J. et al. Mapping the neurodevelopmental predictors of psychopathology. *Mol. Psychiatry* **30**, 478–488 (2025).
26. Cui, Z. & Gong, G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage* **178**, 622–637 (2018).
27. Scheinost, D. et al. Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage* **193**, 35–45 (2019).
28. Varoquaux, G. et al. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *NeuroImage* **145**, 166–179 (2017).
29. Durstewitz, D., Koppe, G. & Meyer-Lindenberg, A. Deep neural networks in psychiatry. *Mol. Psychiatry* **24**, 1583–1598 (2019).
30. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* **77**, 534–540 (2020).
31. Dhamala, E., Yeo, B. T. T. & Holmes, A. J. One size does not fit all: methodological considerations for brain-based predictive modeling in psychiatry. *Biol. Psychiatry* **93**, 717–728 (2023).
32. Koutsouleris, N. & Fusar-Poli, P. From heterogeneity to precision: redefining diagnosis, prognosis, and treatment of mental disorders. *Biol. Psychiatry* **96**, 508–510 (2024).

33. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).
34. Ramduny, J. et al. Representing brain-behavior associations by retaining high-motion minoritized youth. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **11**, 155–170 (2026).
35. Greene, A. S. & Constable, R. T. Clinical promise of brain-phenotype modeling: a review. *JAMA Psychiatry* **80**, 848–854 (2023).
36. Rosenberg, M. D., Casey, B. J. & Holmes, A. J. Prediction complements explanation in understanding the developing brain. *Nat. Commun.* **9**, 589 (2018).
37. Rosenberg, M. D. & Finn, E. S. How to establish robust brain–behavior relationships without thousands of individuals. *Nat. Neurosci.* **25**, 835–837 (2022).
38. Chen, J. et al. Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nat. Commun.* **13**, 2217 (2022).
39. Busch, E. L. et al. Dissociation of reliability, heritability, and predictivity in coarse- and fine-scale functional connectomes during development. *J. Neurosci.* **44**, e0735232023 (2024).
40. Haxby, J. V., Guntupalli, J. S., Nastase, S. A. & Feilong, M. Hyperalignment: modeling shared information encoded in idiosyncratic cortical topographies. *eLife* **9**, e56601 (2020).
41. Dubois, J. & Adolphs, R. Building a science of individual differences from fMRI. *Trends Cogn. Sci.* **20**, 425–443 (2016).
42. Anderson, Z., Gratton, C. & Nusslock, R. The value of hyperalignment to unpack neural heterogeneity in the precision psychiatry movement. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **6**, 935–936 (2021).
43. Gao, S., Mishne, G. & Scheinost, D. Nonlinear manifold learning in functional magnetic resonance imaging uncovers a low-dimensional space of brain dynamics. *Hum. Brain Mapp.* **42**, 4510–4524 (2021).
44. Casanova, R. et al. Embedding functional brain networks in low dimensional spaces using manifoldlearning techniques. *Embedding functional brain networks in low dimensional spaces using manifold learning techniques. Front. Neuroinform* **15**, 740143 (2021).
45. Busch, E. L. et al. Multi-view manifold learning of human brain-state trajectories. *Nat. Comput. Sci.* **3**, 240–253 (2023).
46. Busch, E. L., Conley, M. I. & Baskin-Sommers, A. Manifold learning uncovers nonlinear interactions between the adolescent brain and environment that predict emotional and behavioral problems. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* (2024).
47. Fortunato, C. et al. Nonlinear manifolds underlie neural population activity during behaviour. Preprint at *bioRxiv* https://doi.org/10.1101/2023.07.18.549575 (2024).
48. Lindquist, M. A., Smith, B. B., Kannan, A., Zhao, A. & Caffo, B. Measuring the functioning human brain. *Annu. Rev. Stat. Appl.* **12**, 283–309 (2025).
49. Turk-Browne, N. B. Functional interactions as big data in the human brain. *Science* **342**, 580–584 (2013).
50. Bijsterbosch, J. et al. Challenges and future directions for representations of functional brain organization. *Nat. Neurosci.* **23**, 1484–1495 (2020).
51. Talairach, P. J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging* (Thieme Medical Publishers, 1988).
52. Evans, A. et al. 3D statistical neuroanatomical models from 305 MRI volumes. In *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference* Vol. 3, 1813–1817 (IEEE, 1993).
53. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* **17**, 825–841 (2002).
54. Fischl, B., Sereno, M. I., Tootell, R. B. & Dale, A. M. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* **8**, 272–284 (1999).
55. Burgund, E. D. et al. The feasibility of a common stereotactic space for children and adults in fMRI studies of development. *NeuroImage* **17**, 184–200 (2002).
56. Ghosh, S. S. et al. Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11 years of age. *NeuroImage* **53**, 85–93 (2010).
57. Toga, A. W., Thompson, P. M. & Sowell, E. R. Mapping brain maturation. *Trends Neurosci.* **29**, 148–159 (2006).
58. Raschle, N. et al. Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines. *Ann. N. Y. Acad. Sci.* **1252**, 43–50 (2012).
59. Cox, D. D. & Savoy, R. L. Functional magnetic resonance imaging (fMRI) 'brain reading': detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* **19**, 261–270 (2003).
60. Haxby, J. V. et al. A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* **72**, 404–416 (2011).
61. Himmelberg, M. M., Winawer, J. & Carrasco, M. Linking individual differences in human primary visual cortex to contrast sensitivity around the visual field. *Nat. Commun.* **13**, 3309 (2022).
62. Thirion, B. et al. Dealing with the shortcomings of spatial normalization: multi-subject parcellation of fMRI datasets. *Hum. Brain Mapp.* **27**, 678–693 (2006).
63. Poldrack, R. A. Region of interest analysis for fMRI. *Soc. Cogn. Affect. Neurosci.* **2**, 67–70 (2007).
64. Calhoun, V. D. et al. Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder. *Front. Psychiatry* **2**, 75 (2012).
65. Eickhoff, S. B., Yeo, B. T. T. & Genon, S. Imaging-based parcellations of the human brain. *Nat. Rev. Neurosci.* **19**, 672–686 (2018).
66. Finn, E. S. et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
67. Dubois, J., Galdi, P., Paul, L. K. & Adolphs, R. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Phil. Trans. R. Soc. B* **373**, 20170284 (2018).
68. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
69. Kong, R. et al. Individual-specific areal-level parcellations improve functional connectivity prediction of behavior. *Cereb. Cortex* **31**, 4477–4500 (2021).
70. Cui, Z. et al. Linking individual differences in personalized functional network topography to psychopathology in youth. *Biol. Psychiatry* **92**, 973–983 (2022).
71. Pines, A. R. et al. Dissociable multi-scale patterns of development in personalized brain networks. *Nat. Commun.* **13**, 2647 (2022).
72. Keller, A. S. et al. Personalized functional brain network topography is associated with individual differences in youth cognition. *Nat. Commun.* **14**, 8411 (2023).
73. Fu, Z., Liu, J., Salman, M. S., Sui, J. & Calhoun, V. D. Functional connectivity uniqueness and variability? Linkages with cognitive and psychiatric problems in children. *Nat. Ment. Health* **1**, 956–970 (2023).
74. DeYoung, C. G. et al. Beyond increasing sample sizes: optimizing effect sizes in neuroimaging research on individual differences. *J. Cogn. Neurosci.* **37**, 1023–1034 (2025).
75. Braga, R. M. & Buckner, R. L. Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. *Neuron* **95**, 457–471.e5 (2017).

76. Braga, R. M., Van Dijk, K. R. A., Polimeni, J. R., Eldaief, M. C. & Buckner, R. L. Parallel distributed networks resolved at high resolution reveal close juxtaposition of distinct regions. *J. Neurophysiology* **121**, 1513–1534 (2019).
77. Feilong, M., Nastase, S. A., Guntupalli, J. S. & Haxby, J. V. Reliable individual differences in fine-grained cortical functional architecture. *NeuroImage* **183**, 375–386 (2018).
78. Yoo, K. et al. Multivariate approaches improve the reliability and validity of functional connectivity and prediction of individual behaviors. *NeuroImage* **197**, 212–223 (2019).
79. Badwal, M. W., Bergmann, J., Roth, J. H. R., Doeller, C. F. & Hebart, M. N. The scope and limits of fine-grained image and category information in the ventral visual pathway. *J. Neurosci.* **45**, (2025).
80. Feilong, M., Guntupalli, J. S. & Haxby, J. V. The neural basis of intelligence in fine-grained cortical topographies. *eLife* **10**, e64058 (2021).
81. Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **10**, 424–430 (2006).
82. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
83. Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L. C. & Castellanos, F. X. Inscapes: a movie paradigm to improve compliance in functional magnetic resonance imaging. *NeuroImage* **122**, 222–232 (2015).
84. Guntupalli, J. S., Feilong, M. & Haxby, J. V. A computational model of shared fine-scale structure in the human connectome. *PLoS Comput. Biol.* **14**, e1006120 (2018).
85. Nastase, S. A., Liu, Y.-F., Hillman, H., Norman, K. A. & Hasson, U. Leveraging shared connectivity to aggregate heterogeneous datasets into a common response space. *NeuroImage* **217**, 116865 (2020).
86. Anderson, Z., Turner, J. A., Ashar, Y. K., Calhoun, V. D. & Mittal, V. A. Application of hyperalignment to resting state data in individuals with psychosis reveals systematic changes in functional networks and identifies distinct clinical subgroups. *Apert. Neuro* **4**, 10–52294 (2024).
87. Chen, P.-H. (Cameron) et al. A reduced-dimension fMRI shared response model. In *Adv. Neural Inf. Process. Syst.* Vol. 28 https://proceedings.neurips.cc/paper_files/paper/2015/file/b3967a0e938dc2a6340e258630febd5a-Paper.pdf (Curran Associates, 2015).
88. Mennen, A. C., Nastase, S. A., Yeshurun, Y., Hasson, U. & Norman, K. A. Real-time neurofeedback to alter interpretations of a naturalistic narrative. *Neuroimage Rep.* **2**, 100111 (2022).
89. Finn, E. S. et al. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. *NeuroImage* **215**, 116828 (2020).
90. Sava-Segal, C., Richards, C., Leung, M. & Finn, E. S. Individual differences in neural event segmentation of continuous experiences. *Cereb. Cortex* **33**, 8164–8178 (2023).
91. Botch, T. L. & Finn, E. S. Neural representations of concreteness and concrete concepts are specific to the individual. *J. Neurosci.* **44**, (2024).
92. Esteban, O. et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
93. Guntupalli, J. S. et al. A model of representational spaces in human cortex. *Cereb. Cortex* **26**, 2919–2934 (2016).
94. Busch, E. L. et al. Hybrid hyperalignment: a single high-dimensional model of shared information embedded in cortical patterns of response and functional connectivity. *NeuroImage* **233**, 117975 (2021).
95. Van Dijk, K. R., Sabuncu, M. R. & Buckner, R. L. The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* **59**, 431–438 (2012).
96. Bazeille, T., DuPre, E., Richard, H., Poline, J.-B. & Thirion, B. An empirical evaluation of functional alignment using inter-subject decoding. *NeuroImage* **245**, 118683 (2021).
97. Taschereau-Dumouchel, V. et al. Towards an unconscious neural reinforcement intervention for common fears. *Proc. Natl Acad. Sci. USA* **115**, 3470–3475 (2018).
98. Jiahui, G. et al. Predicting individual face-selective topography using naturalistic stimuli. *NeuroImage* **216**, 116458 (2020).
99. Yates, T. S., Ellis, C. T. & Turk-Browne, N. B. Emergence and organization of adult brain function throughout child development. *NeuroImage* **226**, 117606 (2021).
100. Chang, L. J. et al. Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *Sci. Adv.* **7**, eabf7129 (2021).
101. Sawalha, J. et al. Predicting pediatric anxiety from the temporal pole using neural responses to emotional faces. *Sci. Rep.* **11**, 16723 (2021).
102. Kronberg, G. et al. Shared orbitofrontal dynamics to a drug-themed movie track craving and recovery in heroin addiction. *Brain* **148**, 1778–1788 (2025).
103. Turek, J. S., Ellis, C. T., Skalaban, L. J., Turk-Browne, N. B. & Willke, T. L. Capturing shared and individual information in fMRI data. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 826–830 (IEEE, 2018).
104. Jazayeri, M. & Ostojic, S. Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *Curr. Opin. Neurobiol.* **70**, 113–120 (2021).
105. Bandettini, P. A. et al. The challenge of BWAs: unknown unknowns in feature space and variance. *Med* **3**, 526–531 (2022).
106. Calhoun, V. D., Adalı, T. & Pekar, J. J. A method for comparing group fMRI data using independent component analysis: application to visual, motor and visuomotor tasks. *Magn. Reson. Imaging* **22**, 1181–1191 (2004).
107. Allen, E. A., Erhardt, E. B., Wei, Y., Eichele, T. & Calhoun, V. D. Capturing inter-subject variability with group independent component analysis of fMRI data: a simulation study. *NeuroImage* **59**, 4141–4159 (2012).
108. Chen, Z. S. et al. Modern views of machine learning for precision psychiatry. *Patterns* **3**, 100602 (2022).
109. Parkes, L., Satterthwaite, T. D. & Bassett, D. S. Towards precise resting-state fMRI biomarkers in psychiatry: synthesizing developments in transdiagnostic research, dimensional models of psychopathology, and normative neurodevelopment. *Curr. Opin. Neurobiol.* **65**, 120–128 (2020).
110. Ziegler, G., Dahnke, R., Winkler, A. D. & Gaser, C. Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. *NeuroImage* **82**, 284–294 (2013).
111. Petrican, R. & Fornito, A. Adolescent neurodevelopment and psychopathology: the interplay between adversity exposure and genetic risk for accelerated brain ageing. *Dev. Cogn. Neurosci.* **60**, 101229 (2023).
112. De, A. & Chaudhuri, R. Common population codes produce extremely nonlinear neural manifolds. *Proc. Natl Acad. Sci. USA* **120**, e2305853120 (2023).
113. Perich, M. G., Narain, D. & Gallego, J. A. A neural manifold view of the brain. *Nat. Neurosci.* **28**, 1582–1597 (2025).
114. Salhov, M., Bermanis, A., Wolf, G. & Averbuch, A. Approximately-isometric diffusion maps. *Appl. Comput. Harmon. Anal.* **38**, 399–419 (2015).

115. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
116. Moon, K. R. et al. Manifold learning-based methods for analyzing single-cell RNA-sequencing data. *Curr. Opin. Syst. Biol.* **7**, 36–46 (2018).
117. Coifman, R. R. et al. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc. Natl Acad. Sci. USA* **102**, 7426–7431 (2005).
118. Maaten, L. vd & Hinton, G. Visualizing data using *t*-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
119. Tenenbaum, J. B., Silva, V. D. & Langford, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323 (2000).
120. Moon, K. R. et al. Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* **37**, 1482–1492 (2019).
121. Langs, G. et al. Identifying shared brain networks in individuals by decoupling functional and anatomical variability. *Cereb. Cortex* **26**, 4004–4014 (2016).
122. Gonzalez-Castillo, J. et al. Manifold learning for fMRI time-varying functional connectivity. *Front. Hum. Neurosci.* **17**, (2023).
123. Nguyen, N. D. & Wang, D. Multiview learning for understanding functional multiomics. *PLoS Comput. Biol.* **16**, e1007677 (2020).
124. Conley, M. I., Hernandez, J., Salvati, J. M., Gee, D. G. & Baskin-Sommers, A. The role of perceived threats on mental health, social, and neurocognitive youth outcomes: a multicontextual, person-centered approach. *Dev. Psychopathol.* **35**, 689–710 (2023).
125. Sameroff, A. A unified theory of development: a dialectic integration of nature and nurture. *Child Dev.* **81**, 6–22 (2010).
126. Viding, E., McCrory, E., Baskin-Sommers, A., Brito, S. D. & Frick, P. An 'embedded brain' approach to understanding antisocial behaviour. *Trends Cogn. Sci.* **28**, 159–171 (2024).
127. Astle, D. E., Bassett, D. S. & Viding, E. Understanding divergence: placing developmental neuroscience in its dynamic context. *Neurosci. Biobehav. Rev.* **157**, 105539 (2024).
128. Dick, A. S. et al. Meaningful associations in the adolescent brain cognitive development study. *NeuroImage* **239**, 118262 (2021).
129. Owens, M. M. et al. Recalibrating expectations about effect size: a multi-method survey of effect sizes in the ABCD study. *PLoS ONE* **16**, e0257535 (2021).
130. Taşkin, G. & Crawford, M. M. An out-of-sample extension to manifold learning via meta-modeling. *IEEE Trans. Image Process.* **28**, 5227–5237 (2019).
131. Duque, A. F., Morin, S., Wolf, G. & Moon, K. Extendable and invertible manifold learning with geometry regularized autoencoders. In *Proc. 2020 IEEE International Conference on Big Data (Big Data)* 5027–5036 (IEEE, 2020).
132. Huang, J. et al. Learning shared neural manifolds from multi-subject fMRI data. In *2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP)* 01–06 (IEEE, 2022).
133. Bronfenbrenner, U. & Ceci, S. J. Nature–nuture reconceptualized in developmental perspective: a bioecological model. *Psychol. Rev.* **101**, 568–586 (1994).

Acknowledgements

This work was supported by the National Science Foundation Graduate Research Fellowship Program (grant number 2139841 to E.L.B.) and the National Institutes of Health (grant numbers R61 MH128492 and R01 MH069456 to N.B.T.-B.). We thank D. Gearing for his help with code review.

Author contributions

E.L.B. and A.B.-S. conceived of the idea. E.L.B. and A.B.-S. wrote the initial draft. E.L.B. designed the tutorials. E.L.B., A.B.-S. and N.B.T.-B. provided feedback on the original and revised paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to Erica L. Busch.

Peer review information *Nature Mental Health* thanks the anonymous reviewers for their contribution to the peer review of this work.