Morphometricity: Measuring the Neuroanatomical Signature of a Trait

Mert R. Sabuncu[a,b], Tian Ge[a,c,d], Avram J. Holmes[a,e,f], Jordan W. Smoller[c,d], Randy L. Buckner[a,g], and Bruce Fischl[a,b]; for the Alzheimer's Disease Neuroimaging Initiative (ADNI)*

[a] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129, USA; [b] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02138, USA; [c] Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA; [d] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02138, USA; [e] Department of Psychology, Yale University, New Haven, CT, 06520, USA; [f] Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA; [g] Department of Psychology and Center for Brain Science, Harvard University, Cambridge, MA 02128, USA.

* Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Complex physiological and behavioral traits, including neurological and psychiatric disorders, often associate with distributed anatomical variation. This paper introduces a global metric, called *morphometricity*, as a measure of the anatomical signature of different traits. Morphometricity is defined as the proportion of phenotypic variation that can be explained by macroscopic brain morphology. We estimate morphometricity via a linear mixed effects model that utilizes an *anatomical similarity matrix* computed based on measurements derived from structural brain Magnetic Resonance Imaging (MRI) scans. We examined over 3,800 unique MRI scans from 9 large-scale studies to estimate the morphometricity of a range of phenotypes, including clinical diagnoses such as Alzheimer's disease; and nonclinical traits such as measures of cognition. Our results demonstrate that morphometricity can provide novel insights about the neuroanatomical correlates of a diverse set of traits, revealing associations that might not be detectable through traditional statistical techniques.

Significance

Neuroimaging has largely focused on two goals: mapping associations between neuroanatomical features and phenotypes, and building individual-level prediction models. This paper presents a complementary analytic strategy called *morphometricity* that aims to measure the neuroanatomical signatures of different phenotypes. Inspired by prior work on heritability, we define *morphometricity* as the proportion of phenotypic variation that can be explained by brain morphology, e.g., as captured by structural brain Magnetic Resonance Imaging (MRI). In the dawning era of large-scale datasets comprising traits across a broad phenotypic spectrum, morphometricity will be critical in prioritizing and characterizing behavioral, cognitive, and clinical phenotypes based on their neuroanatomical signatures. Furthermore, the proposed framework will be significant in dissecting the functional, morphological, and molecular underpinnings of different traits.

Introduction

The structural, functional, and molecular properties of the brain support numerous traits spanning the behavioral, cognitive and clinical spectra. Neuroanatomical features are in turn influenced by factors such as age, sex, training, and genetics [1-4]. Neuroimaging allows us to characterize these *bidirectional* associations by revealing variation in brain structure and function across individuals. Conventional methods that we use to probe these associations aim to anatomically map effects, build prediction models, or test hypotheses. Yet, we do not have a standard technique to measure and compare the often spatially distributed and complex patterns of neuroanatomical correlates of different phenotypes. Here we present a novel metric called *morphometricity* that offers this capability.

To date, structural neuroimaging studies have primarily relied on three classes of analytic approaches. The first strategy is hypothesis-driven and utilizes a regression model to examine associations between behavioral traits or clinical conditions and a small number of *a priori* image-derived measurements, such as those restricted to an anatomical region of interest (ROI) [5]. The ROI-based approach provides useful insights about the underlying biology and can be efficient in limited sample size scenarios, but is restricted to the tested hypothesis. The second approach is exploratory and aims to compute maps of associations by conducting brain-wide tests [6], as exemplified in voxel-based morphometry (VBM) [7], or vertex-wise cortical thickness analysis [see e.g., 8]. Such massive univariate analyses can offer a comprehensive picture of the underlying anatomical associations, yet they can also be inefficient in revealing subtle, multivariate patterns of association, because each anatomical location is typically examined in isolation and the burden of multiple testing correction can constrain statistical power. The third class includes multivariate techniques such as Canonical Correlation Analysis (CCA) [9], Partial Least Squares (PLS) [10], Bayesian inference algorithms [11], or other machine learning methods [12,13]. These studies are focused on either discovering the multivariate patterns of association or demonstrating individual-level prediction capabilities, but the biological interpretation of trained multivariate models can be challenging [14]. Furthermore, these methods often suffer from high computational demand, and can be sensitive to implementation details, such as the choice of learning rule, optimization algorithm, and local optima in training.

We present morphometricity analysis as a novel approach to examine the global statistical association between brain morphology and observable traits. Inspired by prior work on trait heritability in population and statistical genetics [15,16], we define morphometricity as the proportion of phenotypic variation that can be explained by brain morphology, e.g., as captured by measurements derived from structural brain Magnetic Resonance Imaging (MRI) scans. Unlike ROI-based or massive univariate association tests, morphometricity analysis is not concerned with specific anatomical structures or the precise anatomic localization of effects. In contrast to the application of machine learning to population data, the primary aim of morphometricity analysis is not to maximize individual-level prediction accuracy, but examine and quantify statistical associations. The proposed strategy thus affords a unique perspective on the biological underpinnings of different phenotypes, and allows us to compare and contrast imaging modalities, types of anatomical measurements, and processing streams.

Morphometricity is grounded in linear mixed effects (LME) modeling, a classical statistical framework that was recently employed in population genetics to quantify the heritability of a trait using genome-wide genetic variants [17-19]. The model relates the variation in brain morphology computed from brain-wide, MRI-derived measurements to the variation in observable traits, and can be fitted using well-established, robust computational tools. In our implementation, we use FreeSurfer [20], a freely available, widely used, and extensively validated brain MRI analysis software package, to automatically process structural MRI scans and obtain a vector of volumetric measurements across subcortical structures and cortical thickness measurements across the entire cortical mantle, which constitute a comprehensive description of the structural neuroanatomy.

We applied the morphometricity analysis to over 3,800 unique brain MRI scans from 9 large-scale studies and computed the morphometricity of clinical conditions including Alzheimer's disease, attention-deficit hyperactivity disorder (ADHD), schizophrenia, autism spectrum disorder, and Parkinson's disease; and nonclinical traits including sex, age, intelligence, education level, and an array of cognitive measures. Our results demonstrate that morphometricity analysis promises to offer a unique perspective on the relationship between brain anatomy, and behavioral, cognitive and physiological traits.

Results

Overview of the Model. The proposed morphometricity analysis is based on the following linear mixed effects (LME) model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{a} + \boldsymbol{\epsilon}, \qquad (1)$$

where $\boldsymbol{y}$ is an N-dimensional column vector of a quantitative phenotype with N being the sample size (i.e., number of subjects), $\boldsymbol{X}$ is the (design) matrix of confounding variables (sometimes called covariates or nuisance variables) such as age and sex, $\boldsymbol{\beta}$ is the (fixed effect) coefficient vector, $\boldsymbol{a} \sim \mathrm{N}(\boldsymbol{0}, \sigma_a^2 \boldsymbol{K}_a)$ is an N-dimensional random effect vector drawn from a zero-mean multivariate Gaussian distribution with a covariance matrix that is equal to the scaled anatomic similarity matrix (ASM) $\boldsymbol{K}_a$, and the elements of the noise vector $\boldsymbol{\epsilon}$ are assumed to be drawn from independent and zero-mean Gaussian distributions with homogeneous variance σ_e^2. The ASM encodes global morphological resemblance between pairs of individuals in the sample, and in principle can be any non-negative definite matrix with its diagonal elements constrained to be equal to 1 on average, and σ_a^2 can thus be interpreted as the total variance captured by the ASM. In this paper, we considered two types of intuitive and widely used metrics that quantify the similarity of volumetric and cortical thickness measurements extracted from structural brain MRI scans between pairs of individuals: (1) a linear similarity metric (i.e., inner product between normalized imaging measurements); and (2) a nonlinear Gaussian-type similarity metric (see Methods). Using a model selection approach, we found that the Gaussian similarity metric provided consistently better description of the data across the traits we studied (see Methods and Supplementary Table S1). Therefore, all reported morphometricity estimates were based on the Gaussian metric.

Formally, we define morphometricity based on the LME model of Equation (1) as:

$$m^2 \doteq \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2} = \frac{\sigma_a^2}{\sigma_y^2}, \qquad (2)$$

where σ_y^2 is the phenotypic variance. Morphometricity is thus the proportion of phenotypic variation that can be explained by brain morphology, the variation of which is captured by the ASM. Estimates of morphometricity can be computed by plugging the restricted maximum likelihood (ReML) estimates [21,22] of the variance components, σ_a^2 and σ_e^2, into Equation (2).

We extend this definition to the case-control design, i.e., binary disease traits (affected vs. unaffected), using the classical liability-threshold model, which is widely employed in population and statistical genetic studies [23,24]. The model assumes that the underlying disease liability (which is a quantitative variable) follows a Gaussian distribution and individuals are cases (affected) if their liability exceeds a threshold. The morphometricity estimate m^2 for a disease trait on the *observed scale*, obtained by fitting the model of Equation (1) to the binary phenotype data, can be easily transformed to the *liability scale* [24,25]:

$$m_l^2 = m^2 \frac{K(1-K)}{\varphi(t)^2} \frac{K(1-K)}{P(1-P)}, \quad (3)$$

where m_l^2 is the morphometricity on the liability scale, K is the prevalence of the disease in the general population (i.e., the proportion of the population having the disease), P is the prevalence of the disease in the study sample, $t = \Phi^{-1}(1-K)$ is the liability threshold, Φ is the standard Gaussian cumulative distribution function, and φ is the standard Gaussian density function. In real disease studies, cases are often considerably oversampled relative to their population prevalence (known as non-random ascertainment), in which case P is larger than K. Transforming morphometricity estimates from the observed scale to the liability scale makes them independent of population and sample prevalence, and thus comparable across different diseases.

Overview of the Data. We analyzed brain MRI scans and trait data from over 3,800 unique individuals spanning 9 large-scale studies: the Harvard/Massachusetts General Hospital Brain Genomic Superstruct Project (GSP) [26], the Human Connectome Project (HCP) [27], the Alzheimer's Disease Neuroimaging Initiative (ADNI) [28], the Attention-Deficit Hyperactivity Disorder (ADHD 200) sample [29], the Open Access Series of Imaging Studies (OASIS) cross-sectional sample [30], the Center for Biomedical Research Excellence (COBRE) schizophrenia sample [31], the MIND Clinical Imaging Consortium (MCIC) schizophrenia sample [32], the Autism Brain Imaging Data Exchange (ABIDE) [33], and the Parkinson Progression Marker Initiative (PPMI) [34]. See the Methods section for further details on the datasets. The traits of interest were grouped into three categories: clinical diagnoses, general nonclinical traits and measures of cognition.

Morphometricity of Clinical Diagnoses. The clinical traits we examined included Alzheimer's disease, attention-deficit hyperactivity disorder (ADHD), schizophrenia, autism spectrum disorder, and

Parkinson's disease. Table 1 lists the characteristics of the corresponding samples, along with morphometricity estimates (on the liability scale) and assumed population prevalence values [35-41]. Figure 1 shows the estimated morphometricity values on the liability scale. Our analyses revealed that Alzheimer's disease is substantially morphometric (with a 95% confidence interval of [0.94-1.00]), suggesting that this clinical condition is associated with a substantial anatomical signature. On the other hand, ADHD, schizophrenia and autism showed moderate morphometricity values, all greater than 0.35. Finally, we found that Parkinson's disease was modestly morphometric, with an estimated liability-scale value of 0.20. All examined clinical conditions were statistically significantly associated with whole-brain macroscopic morphology, i.e., the estimated morphometricity values were significantly larger than zero (all p-values < 0.005). Supplementary Table S2 lists point estimates of morphometricity and their standard errors computed via jackknife resampling [42]. These results are in strong agreement with the parametric estimates in Table 1.

We had access to two independent samples that allowed us to replicate our morphometricity estimates of Alzheimer's disease (OASIS) and schizophrenia (COBRE). We observed that there was strong agreement between the estimates from independent samples (Figure 1). Supplementary Table S3 provides further data about these replication analyses.

Morphometricity of General, Nonclinical Traits. The nonclinical traits we examined were age, general intelligence (IQ), sex, and education level. Table 2 lists the characteristics of the samples used in the primary analyses, along with the estimates of morphometricity. These results revealed that all the examined general traits are significantly and substantially morphometric (Figure 2); all morphometricity point estimates were greater than 0.8, and all p-values < 1e-8. Supplementary Table S2 lists morphometricity estimates and their standard errors computed via jackknife resampling. These results and the parametric estimates in Table 1 are virtually identical.

We had access to independent replication samples for all the general nonclinical traits we examined. It can be seen in Figure 2 that the replication analyses revealed remarkably consistent morphometricity estimates. Supplementary Table S4 provides further data about these replication analyses.

Contrasting with ROI-based Association Analyses. The most common analytic strategy in today's neuroimaging studies involves examining associations between traits and measurements from regions of interest (ROIs). Our goal in this analysis was to contrast the proposed whole-brain morphometricity analysis with such ROI-based techniques. We restricted our analysis to the six phenotypes (age, IQ, sex, education, Alzheimer's disease, and schizophrenia), for which we had two independent datasets. We then used one of the samples (the replication sample in the analyses above) for the discovery of the most significantly associated ROI with the trait, and the other sample (the primary sample in the analyses above) to quantify the strength and magnitude of association. Supplementary Table S5 lists the structures that exhibited the strongest association with the traits in the discovery analysis.

Figure 3 visualizes the magnitude of association between the ROI-based measurements and phenotypic variation, assessed using the same LME modeling framework of whole-brain morphometricity. Here, we replaced the global ASM with one computed based on ROI measurements (see Methods for further details). It can be seen that the proportion of variance explained by ROI-based measurements was consistently lower than whole-brain morphometricity estimates (with general intelligence exhibiting the smallest discrepancy). Most notably, for education and schizophrenia, ROI-based associations were much weaker (both in magnitude and statistical significance) than whole-brain associations (Figure 3 and Supplementary Table S5). In fact, education and schizophrenia did not exhibit a statistically significant correlation with individual ROI-based measurements, while whole-brain morphometricity analyses revealed significant associations.

Morphometricity Analysis of Cognitive Measures. We used the most recent release of the Human Connectome Project (HCP) data (downloaded on December 15, 2015) to compute morphometricity estimates for an array of cognitive measures. Our primary analysis relied on 190 non-twin subjects of non-Hispanic European ancestry (28.9+/-3.8 years, 47.3% female), drawn from separate families (i.e., there were no siblings in this sample). Figure 4 shows the morphometricity estimates computed for variables that measure sustained attention, non-verbal and verbal episodic memory, working memory, executive function, delay discounting, language (vocabulary comprehension and reading decoding), spatial orientation, processing speed, fluid intelligence, and self-regulation (impulsivity). We conducted a secondary (replication) analysis on 208 non-Hispanic white twins (29.4+/-3.2 years,

61.4% female), with one twin drawn from each family, and thus there were no siblings in this replication sample. Figure 4 shows the results obtained from this secondary analysis as well.

Our results demonstrated that all examined variables, except for the measure of self-regulation, are statistically significantly morphometric. We note further that there was an important amount of variability in the degree of morphometricity across cognitive measures. This variation was remarkably consistent between the primary and secondary analyses. Measures of attention, cognitive flexibility, working memory, verbal episodic memory, and inhibition were substantially morphometric (with estimates greater than 0.80 in both primary and secondary analyses). Measures of language, non-verbal episodic memory, spatial orientation, processing speed, and fluid intelligence were moderately morphometric (with point estimates greater than 0.50).

Discussion

Morphometricity: A Novel Metric to Quantify Whole-brain Associations with a Trait. In this paper, we introduced a novel technique for analyzing the neuroanatomical underpinnings of various clinical, physiological, and behavioral traits using large-scale neuroimaging data. In contrast with association testing techniques widely used in today's neuroimaging studies, our approach does not focus on *a priori* ROIs or conduct independent (massive univariate) interrogations at each candidate region or voxel. Instead, morphometricity is a global quantification of the whole-brain anatomical signature of a trait.

While the proposed approach is intimately related to image-based multivariate prediction performance, there are two characteristics of morphometricity that make it different from the application of machine learning. Firstly, the metric does not require cross-validation, which is often the technique utilized in machine learning to gauge prediction accuracy. Cross-validation is usually computationally demanding, and relies on the unbiased setting of model parameters (which might be achieved via a nested cross validation strategy), and repeated and balanced partitioning of data into train and test sets. In contrast, the proposed LME-based approach exploits the entire dataset to fit the model and estimate the unknown variance component parameters, and in turn morphometricity,

in an unbiased fashion. Secondly, morphometricity is a classical statistical measure of explained variance and is therefore familiar to interpret.

Morphometricity has many parallels to heritability in genetics [15,16]. Both concepts, statistical in nature, are about phenotypic variation and the proportion of variance explained. Thus, the interpretation has to be carried out within a probabilistic framework, is limited by the studied population, depends on the technique used to quantify the trait, and can be confounded by unmeasured variables acting through unknown mechanisms. The biases due to confounds such as (cryptic) relatedness between subjects or population admixture are well-studied in heritability analysis. As we elaborate below, morphometricity alone cannot be used to infer causal relationships, but has to be followed up with further studies that will home in on potential mechanisms. The core difference between morphometricity and heritability is the direction of association. In heritability, this direction is known and fixed, because there is no known biological mechanism that would allow the phenotype to alter the genotype. In morphometricity, however, the directionality can go either way and has to be dissected with further biological studies.

Traits Can Be Morphometric to Different Degrees. Virtually all traits we examined in this study were significantly morphometric. However, our analyses also revealed interesting variation in the whole-brain anatomical signature of different traits. Certain phenotypes, such as Alzheimer's disease, age, and (maybe surprisingly) general intelligence (IQ) were substantially morphometric (with estimates exceeding 0.90), while other measures, such as non-verbal episodic memory, spatial orientation, processing speed, and fluid intelligence exhibited moderate morphometricity. Furthermore, the psychiatric disorders we examined (schizophrenia, ADHD and autism) were all moderately morphometric, unequivocally pointing to a neuroanatomical substrate for these clinical conditions. The proposed morphometricity analysis is the first coherent framework that enables us to directly quantify and compare the morphological signatures of such diverse sets of traits.

The traits we presented in this study have been examined extensively in prior structural neuroimaging studies to reveal morphological correlates. While many of these studies relied on regional or voxel-level association tests that are conducted at each location independently, there is growing evidence that multiple brain regions are implicated in complex, multivariate relationships

with many common phenotypes. For example, patterns of atrophy in neurodegenerative conditions such as Alzheimer's disease have been shown to spread through large-scale, distributed brain networks that can be circumscribed based on resting-state activity [43]. Furthermore, developmental mechanisms such as neuronal migration, synapse formation, myelination, and synaptic pruning follow predictable and robust spatiotemporal patterns, which are likely associated with behavioral traits such as intelligence and disrupted in psychiatric disorders such as schizophrenia [44,45]. Our data from ROI-based analyses support the premise of analyzing brain-wide patterns, rather than isolated regions, for associations between neuroanatomical features and behavioral or clinical phenotypes.

Morphometricity Estimates Are Consistent Across Independent Samples. For most of the traits we examined in this study, we replicated our analyses on independent data. All point estimates fell within 95% confidence intervals of the estimates computed on the corresponding independent data. These results suggested that the presented morphometricity estimates are consistent across different samples. We present these results with a cautionary note however. As we emphasized above, morphometricity is a statistical metric that depends on the studied population and the measurement of the trait. Thus, variations in the patient composition for example or changes in diagnostic criteria will inevitably lead to different estimates of morphometricity. In our replication analyses, these factors seemed to play only a minor role.

Whole-brain Morphometricity Analyses Can Be More Powerful than ROI-based Analyses. We present morphometricity analysis as an alternative to the classical region-based interrogation conducted in neuroimaging, which is often focused on discovering or characterizing biomarkers and mapping biological effects. The central challenge in region-based approaches is that we need to either confine our analyses to *a priori* ROIs, or exhaust statistical power by probing a large number of candidate regions. In our experiments, we conducted a direct comparison between whole-brain morphometricity analysis and an ROI-based approach. To identify trait-specific ROIs, a discovery analysis was run on independent samples of each trait. The associations between the identified ROIs and traits were then tested in non-overlapping samples. Identifying the most associated ROI and estimating the magnitude of association in independent samples avoided the issue of circular analysis [46,47], and produced unbiased morphometricity estimates for individual ROIs. Our results demonstrated that whole-brain morphology consistently explained more of the phenotypic variation

than single ROIs. Furthermore, morphometricity analysis could reveal associations that were not detectable when focused on isolated regions. For example, education and schizophrenia were found to be not significantly associated with volumetric/thickness measurements of any of the individual ROIs; yet, both traits were moderately and significantly morphometric in whole-brain analyses, which indicates that they may have spatially distributed neuroanatomical signatures that cannot be captured by individual ROIs. In addition, whole-brain morphometricity analysis offers the capability of capturing interactions between brain regions and thus can be more powerful than analyzing each ROI independently.

Potential Limitations and Drawbacks of Morphometricity. Morphometricity is a statistical metric and assumes a particular, linear model of the relationship between variables. One critical component of the model is the Anatomical Similarity Matrix (ASM), which captures the covariance structure of the random effect that accounts for the morphological variation in the sample. In this work, we considered a linear metric and a nonlinear Gaussian-type metric to quantify the similarity of volumetric/thickness measurements between pairs of subjects, and employed a model selection technique to find the metric that better describes the data. Our analyses suggest that the Gaussian metric is consistently better than the linear metric across the traits we studied, and captures a significant portion of relevant inter-subject variation under different conditions. However, we have not attempted to exhaustively explore other types of similarity metrics, which may emphasize different aspects of the data and produce different results. Alternatively, the ASM can be built using a bottom-up approach and expressed as a combination of elementary matrices, and the parameters of this combination can be treated as unknown variables. By increasing the unknowns in the model, however, this approach will likely reduce statistical power.

Another important point to consider is that ASM should reflect brain-wide global morphology and cannot be optimized for a specific trait. This latter observation is critical to be able to objectively compare morphometricity estimates across different traits. Our global definition of ASM and morphometricity, however, constrains the interpretation of the results. Certain traits with very dramatic yet focal effects, might not yield large morphometricity estimates, since the proposed model is insensitive to localized effects.

Finally, as in any association testing framework, the interpretation of morphometricity analysis results should be done very carefully and consider confounding mechanisms. This is well understood in the context of heritability, where non-genetic (e.g., dietary, cultural, socioeconomic) influences that vary across racial groups or families can confound genetic analyses. Similar drawbacks apply to morphometricity. For example, siblings might have similar brain morphologies and phenotypic expressions, yet there might not be any causal link between brain morphology and the phenotype. Therefore, we recommend running morphometricity analyses on a set of unrelated subjects. Furthermore, we advise constraining the analysis (if possible) to a homogeneous sample of uniform ancestry, and explicitly controlling for other potential confounding factors, such as age, sex, and scan site.

Potential Uses and Extensions of Morphometricity. Morphometricity analysis can be used to prioritize imaging modalities, acquisition parameters, and processing pipelines. For example, there are a growing number of software packages that allow us to automatically extract numerous structural measurements from brain images. Different constructions of feature vectors and different metrics that quantify the similarity of imaging features between individuals will result in distinct ASM's, which can be compared with the model selection framework employed in this study. Therefore, morphometricity analysis offers a way to quantitatively and objectively identify the imaging features and inter-subject similarity metric that best describe the trait-relevant aspects of whole-brain morphology.

Alternatively, one can imagine estimating the functional, structural, connectomic, and molecular signatures of a trait within a single statistical model, where each of these components is represented with a random effect and the corresponding similarity matrix computed from a relevant modality. Similarly, one can partition the phenotypic variation into contributions from, for instance, cortical and subcortical features, or different large-scale brain networks. This strategy might offer novel insights about the neural correlates of certain phenotypes by integrating multiple modalities and/or modeling spatial heterogeneity in a unified analytic framework. This novel perspective might also allow us to quantify the complementary information contained in different imaging modalities and spatial locations.

As longitudinal imaging studies continue to grow, it will be interesting to extend morphometricity to examine the relationships between temporal dynamics in brain morphology (e.g., global atrophy rates) and clinical or behavioral traits. We envision employing the linear mixed effects strategy to model longitudinal data [48,49], and we will define the morphometricity of longitudinal changes in phenotypes within this framework.

Finally, the proposed framework can also offer a novel perspective on examining relationships between different phenotypes. We plan to extend morphometricity, which is essentially the degree of association between global brain morphology and a phenotype, to quantify the "morphological correlation" between phenotypes. This will be analogous to genetic correlation analysis [50], which quantifies the genetic overlap between traits. We believe that morphological correlation will be an invaluable tool to examine the complex biological relationships between the various dimensions of human behavior and will inform basic and translational research into exploring and redefining the landscape of brain diseases.

Methods

The Imaging Measurements. In this study, we used the extensively studied FreeSurfer-derived measurements to describe the whole-brain morphology. The imaging measurements included volumes of non-cortical structures [51] (left and right cerebral white matter, lateral ventricle, inferior lateral ventricle, cerebellum white matter, cerebellum cortex, thalamus proper, caudate, putamen, pallidum, hippocampus, amygdala, and the 3rd and 4th ventricles) and thickness measurements of cortical regions [52] (left and right superior frontal, rostral middle frontal, caudal middle frontal, pars opercularis, pars triangularis, pars orbitalis, lateral orbitofrontal, medial orbitofrontal, precentral, paracentral, frontal pole, superior parietal, inferior parietal, supra marginal, post central, precuneus, superior temporal, middle temporal, inferior temporal, banks of the superior temporal sulcus, fusiform, transverse temporal, entorhinal, temporal pole, parahippocampal, lateral occipital, lingual, cuneus, pericalcarine, rostral anterior frontal, caudal anterior frontal, posterior parietal, isthmus parietal, and insula).

The Anatomical Similarity Matrix. The anatomical similarity matrix (ASM) plays a central role in the proposed morphometricity analysis. The ASM is an $N \times N$ symmetric matrix, where N is the number of subjects in the analyzed sample. Entries in the ASM quantify the pairwise global similarity between the brain morphologies of two individuals. In principle, the ASM can be any non-negative definite matrix with its diagonal elements constrained to be equal to 1 on average. In this study, we considered two widely used similarity metrics (linear and Gaussian) to construct the ASM. Assume that v_{ik} denotes the k-th imaging measurements from subject i, M is the total number of measurements, and s_k is the sample standard deviation of the k-th measurement. The first ASM we considered uses a *linear kernel*. Thus, the similarity is measured as the linear *correlation* between pairs of imaging vectors and the (i, j)-th entry is computed as:

$$\frac{1}{M}\sum_{k}\frac{v_{ik}v_{jk}}{s_k^2}. \quad (4)$$

This is equivalent to modeling the random effect $\boldsymbol{a}$ in Equation (1) as a linear combination of the imaging features: $\boldsymbol{a} = \boldsymbol{Z}\boldsymbol{u}$, where $\boldsymbol{Z}$ is an $N \times M$ matrix comprising the standardized imaging measurements (i.e., the (i, k)-th entry of $\boldsymbol{Z}$ is v_{ik}/s_k), and $\boldsymbol{u}$ is an $M \times 1$ random vector distributed as $\boldsymbol{u} \sim \mathrm{N}\left(\boldsymbol{0}, \frac{\sigma_a^2}{M}\boldsymbol{I}\right)$, i.e., each imaging measurement is associated with an independent and normally

distributed effect size. The covariance of $\boldsymbol{a}$ can then be computed as $\sigma_a^2 \cdot \boldsymbol{Z}\boldsymbol{Z}^T/M$, i.e., $\boldsymbol{Z}\boldsymbol{Z}^T/M$ is the ASM with its entries explicitly stated in Equation (4).

The second ASM we considered was computed using *a Gaussian* kernel on standardized imaging features, with the (i, j)-th entry defined as:

$$\exp\left(-\sum_k \frac{(v_{ik} - v_{jk})^2}{M s_k^2}\right). \quad (5)$$

Note that each ASM definition corresponds to different models of trait-relevant variation. For example, the Gaussian kernel can capture nonlinear and multivariate associations between brain morphology and traits. One can further decide to weigh different features differently, based on for example, some a priori information about trait relevance. Each of these ASM choices will correspond to a particular semi-parametric regression model, where the morphological association is captured with a function that belongs to a specific space of functions induced by the utilized kernel [53]. Below, we describe an empirical strategy to choose the most appropriate ASM model from a selection of candidates.

In the presented study, for a given similarity metric, we computed an ASM for the cortical thickness measurements and an ASM for the head-size-normalized volumes of non-cortical structures (i.e., divided by total intracranial volume estimates). The global ASM was then computed as the average of the cortical and non-cortical ASM's.

Model Selection. To select the ASM that can best describe the data, we used a model selection technique derived for LME models and proposed in [53]. Specifically, for a given ASM $\boldsymbol{K}_a$, if we denote $\widehat{\boldsymbol{V}} = \hat{\sigma}_a^2 \boldsymbol{K}_a + \hat{\sigma}_e^2 \boldsymbol{I}$ as the ReML estimate of the covariance of $\boldsymbol{y}$, $\widehat{\boldsymbol{P}} = \widehat{\boldsymbol{V}}^{-1} - \widehat{\boldsymbol{V}}^{-1}\boldsymbol{X}(\boldsymbol{X}\widehat{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}\widehat{\boldsymbol{V}}^{-1}$, and define $\boldsymbol{S} = \boldsymbol{I} - \hat{\sigma}_e^2 \widehat{\boldsymbol{P}}$, it can be shown that $\mathrm{tr}(\boldsymbol{S})$, i.e., the trace of the matrix $\boldsymbol{S}$, is a measure of model complexity. Liu et al. [53] thus proposed the following Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in the LME modeling framework:

$$\mathrm{AIC} = N\log(RSS) + 2\mathrm{tr}(\boldsymbol{S}), \quad \mathrm{BIC} = N\log(RSS) + \log(N) \cdot \mathrm{tr}(\boldsymbol{S}), \quad (6)$$

where $RSS = (\boldsymbol{y} - \widehat{\boldsymbol{y}})^T(\boldsymbol{y} - \widehat{\boldsymbol{y}}) = \hat{\sigma}_e^2(N - \mathrm{tr}(\boldsymbol{S}))$ is the residual sum of squares (RSS) of the LME model. Both AIC and BIC reward the goodness-of-fit (the first term) of the model and penalize

complex models (the second term) to avoid overfitting. BIC has a larger penalty term than AIC for large N and thus favors simpler models. We selected the ASM that gave smaller AIC/BIC values.

The Analysis Pipeline. We first ran FreeSurfer on all structural brain MRI scans available in each of the analyzed studies, and dropped the subjects that FreeSurfer failed to complete successfully. Next, we conducted automatic quality control on the measurements computed by FreeSurfer by identifying outliers – if > 25% of the utilized morphometric variables exhibited values that were more than 2 standard deviations away from the population mean, we deemed that subject an outlier and discarded it. For the remaining subjects, we computed pairwise similarity measures based on the linear or Gaussian kernel described above. If there were two subjects that exhibited a cortical and non-cortical similarity measure greater than 0.95, we dropped one of those subjects, accounting for the possibility that this was a duplicate case or a closely related individual. Finally, we included sex and age as covariates in all our analyses (unless sex or age was the trait of interest, in which case that variable was not included). We further introduced dummy variables that indicated site IDs when analyzing multi-site data (except for ADNI, where there were a large number of sites, yet the imaging parameters were carefully calibrated across sites [28]).

Given individual-level data and the ASM, we fit the model of Equation (1) to estimate the variance component parameters via the Restricted Maximum Likelihood (ReML) algorithm [21,22]. In Matlab, we implemented an efficient Fisher scoring method to iteratively maximize the restricted likelihood of the model. The standard error of the variance component estimates can be derived using the inverse of the Fisher information matrix when the algorithm converges. Empirically we confirmed the parametric estimates using jackknife resampling (see Supplementary Table S2). Significance of the morphometricity estimate was obtained via a likelihood ratio test, comparing LME models with and without the random effects. Because the null hypothesis ($\sigma_a^2 = 0$) lies on the boundary of the parameter space, the likelihood ratio test statistic follows a half-half mixture of χ_0^2 (a chi-square distribution with all probability mass at zero) and χ_1^2 (a chi-square distribution with one degree of freedom) [54]. Implementations of the developed morphometricity tools are available at: http://people.csail.mit.edu/msabuncu/morphometricity.

ROI-based Morphometricity Analysis. In order to compare the proposed whole-brain morphometricity analysis to more conventional ROI-based analyses, we implemented an ROI-based adaptation of Equation (1). Here, instead of computing the ASM using an array of variables that span the brain, we computed ROI-based ASMs only using a single ROI biomarker – the variable that had the strongest association with the trait of interest. To identify the ROI biomarker of each trait, we conducted an independent discovery analysis on a non-overlapping sample, examining the association between each of the candidate imaging variables and the trait in a regression analysis, while appropriately controlling for age, sex, and site. The imaging variable that exhibited the smallest p-value was then identified as the ROI biomarker for the trait of interest and used in the morphometricity analysis. We note that for the ROI-discovery analyses we utilized the replication (secondary) samples of the whole-brain morphometricity analyses. This way, we computed the ROI-based morphometricity results using the primary samples.

The Data. We employed baseline brain MRI scans (T1-weighted acquired on 1.5T machines), clinical diagnosis, and demographic variables from phase 1 of the ADNI [28]. In the ADHD 200 sample [29], cases were defined as those with evidence of non-typical development and an ADHD-Combined diagnosis, as per the published phenotypic key [http://fcon_1000.projects.nitrc.org/indi/adhd200/general/ADHD 200_PhenotypicKey.pdf]. In the cross-sectional OASIS sample [30], subjects (of 60 years or older) with a clinical dementia rating (CDR) greater than 0 were classified as having dementia. Elderly subjects with a 0 CDR were classified as healthy controls. For the 20 control subjects with repeat scans, we only used the data from the first imaging session. In the COBRE sample [31], schizophrenia subjects were identified according to the phenotype file [http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html]. The MCIC data was compiled from a shared repository of multi-site brain imaging data collected for the clinical investigation of schizophrenia [32]. The ABIDE analyses [33] were conducted on subjects who were older than 10 years, and cases were defined as those having a non-zero diagnostic group entry in the phenotype table [http://fcon_1000.projects.nitrc.org/indi/abide/]. In the PPMI analyses [34], cases were determined to be those diagnosed with Parkinson's Disease at baseline and controls were those who were clinically healthy and not prodromal, again at baseline [http://www.ppmi-info.org/access-data-specimens/download-data/].

All general trait analyses were conducted on healthy control samples (see criteria of relevant study). Both the OASIS and GSP samples cover a substantial portion of the adult life span and thus were used to estimate the morphometricity of age. All GSP analyses were constrained to unrelated, healthy controls of non-Hispanic European ancestry, with high-quality structural brain MRI scans acquired on a 12-channel coil [26]. In the general intelligence (IQ) analyses, we utilized the Wechsler Abbreviated Scale of Intelligence as the phenotype (both in the ADHD 200 and ABIDE samples). For the morphometricity analysis of sex, we created subsamples that were gender-balanced (50% female) and age-matched between sexes. In the PPMI data, education was measured in years (minimum 9 and maximum 24), whereas in the OASIS sample, education levels were encoded as; 1: less than high school graduate, 2: high school graduate, 3: some college, 4: college graduate, and 5: beyond college.

In the cognitive measure analyses, we employed the demographic and behavioral measures reported in the 'open access' and 'restricted' subject information spreadsheets available from the HCP database website [http://www.humanconnectome.org/data]. The HCP collected a range of well-validated and reliable behavioral measures, including those from the NIH Toolbox Assessment of Neurological and Behavioral Function, and several additional measures to assess domains not covered by the NIH Toolbox. For more information on the rationale behind the development of the behavioral batteries used in HCP, see [55].

All MRI scans from ADNI, OASIS, ADHD 200, MCIC, COBRE, PPMI, and HCP were processed with FreeSurfer version 5.3. The GSP MRI scans were processed with FreeSurfer version 4.5.

Acknowledgements

Support for this research was provided in part by the National Institute for Biomedical Imaging and Bioengineering (P41EB015896, R01EB006758, R21EB018907, R01EB019956), the National Institute on Aging (5R01AG008122, R01AG016495), the National Institute for Neurological Disorders and Stroke (R01NS0525851, R21NS072652, R01NS070963, R01NS083534, 5U01NS086625), and was made possible by the resources provided by Shared Instrumentation Grants 1S10RR023401, 1S10RR019307, and 1S10RR023043. Additional support was provided by the NIH Blueprint for Neuroscience Research (5U01-MH093765), part of the multi-institutional Human Connectome Project.

Data were provided in part by the Brain Genomics Superstruct Project (GSP) of Harvard University and MGH, with support from the Center for Brain Science Neuroinformatics Research Group, Athinoula A. Martinos Center for Biomedical Imaging, Center for Human Genetic Research, and Stanley Center for Psychiatric Research. 20 individual investigators at Harvard and MGH generously contributed data to the overall project.

This research was also funded in part by NIH grants R01NS083534, R01NS070963, and 1K25EB013649-01 and 1R21AG050122-01A1 (to MRS); K01MH099232 (to AJH); K24MH094614 and R01MH101486 (to JWS); an MGH ECOR Tosteson Postdoctoral Fellowship Award (to TG). JWS is a Tepper Family MGH Research Scholar. BF has a financial interest in CorticoMetrics, a company whose medical pursuits focus on brain imaging and measurement technologies. BF's interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies.

Tables and Figures

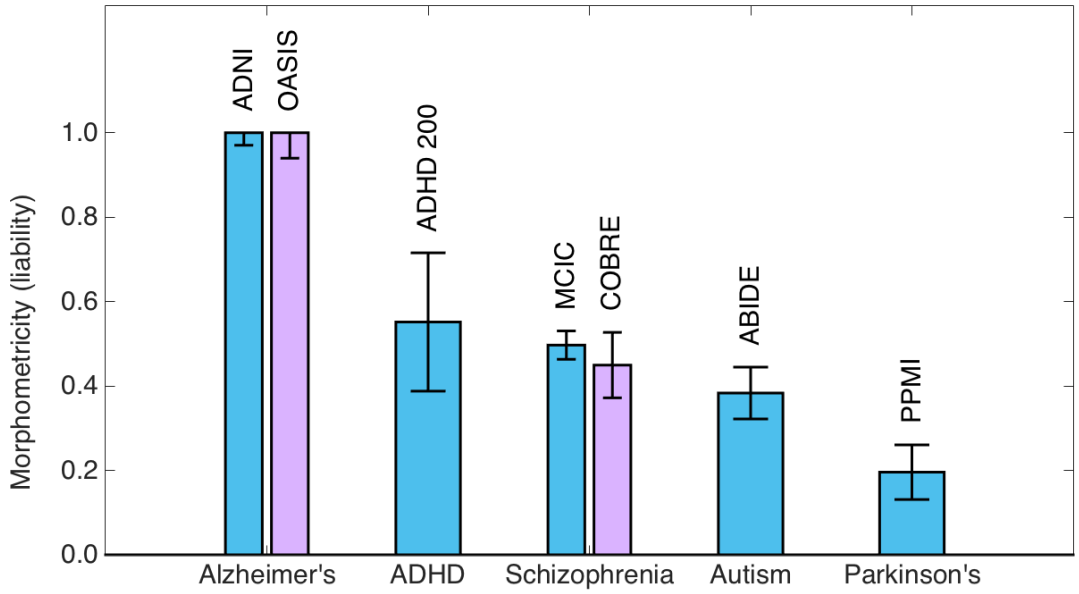

Figure 1: Morphometricity estimates of various diseases (on the liability scale) computed using the Gaussian anatomical similarity matrix (ASM) of Equation (5). Each bar is annotated with study names used to compute these estimates. For Alzheimer's disease and schizophrenia, we had independent samples used to compute replication estimates (purple bars). Error bars indicate standard error of the estimates.

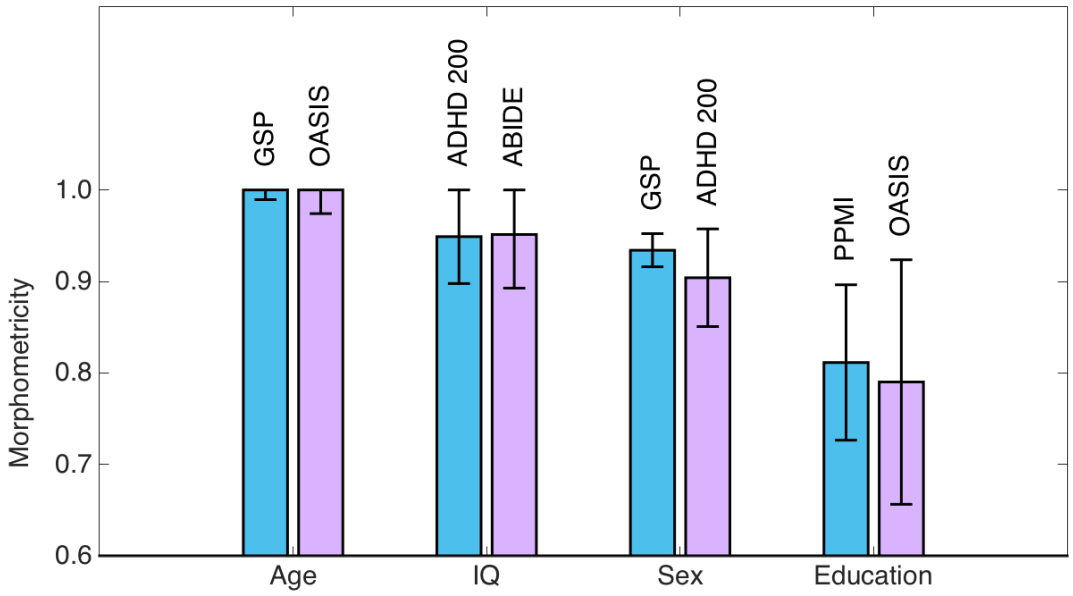

Figure 2: Morphometricity estimates of general nonclinical traits computed using the Gaussian anatomical similarity matrix (ASM) of Equation (5). IQ denotes general intelligence. Each bar is annotated with study names used to compute these estimates. Blue bars correspond to results from the primary analyses, whereas purple bars correspond to independent replication analyses. Error bars indicate standard error of the estimates.

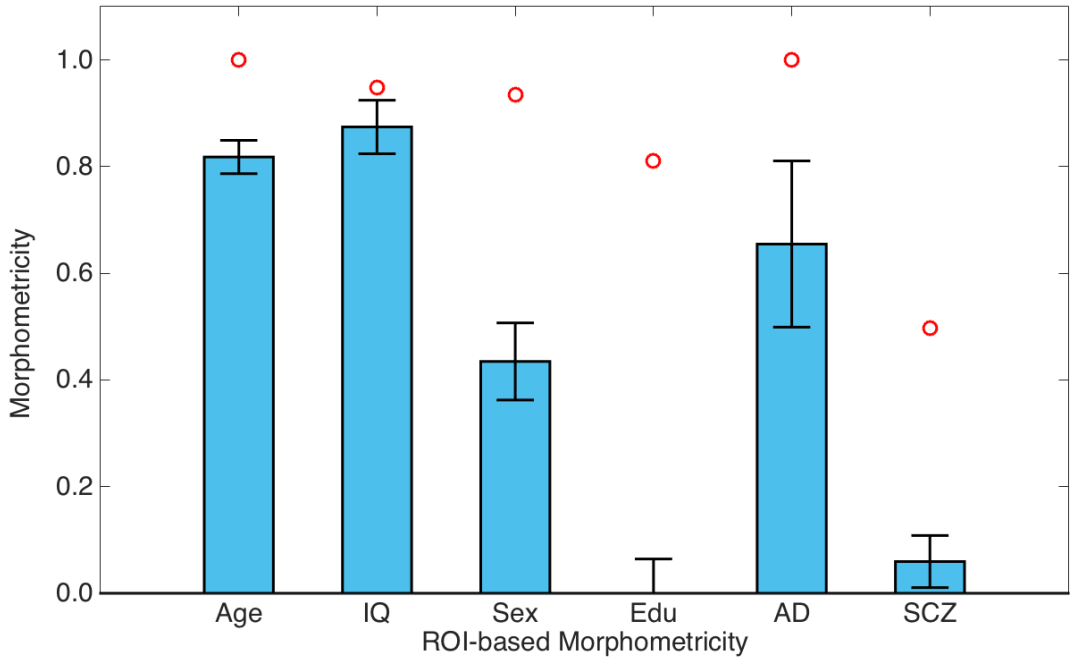

Figure 3: ROI-based morphometricity estimates of general nonclinical traits (age, intelligence, sex, and education), Alzheimer's disease (AD), and schizophrenia (SCZ). For AD and SCZ, morphometricity estimates have been transformed to the liability scale. Red circles denote whole-brain morphometricity estimates for each trait. Error bars indicate standard error of the estimates.

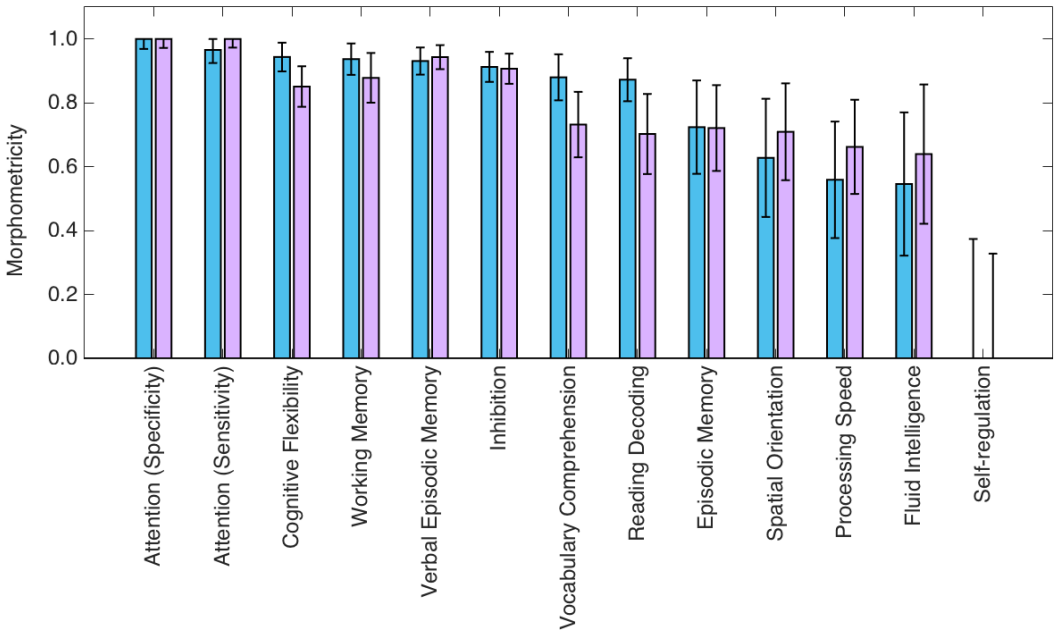

Figure 4: Morphometricity estimates of various measures of cognition computed on data from the Human Connectome Project (HCP) and using the Gaussian anatomical similarity matrix (ASM). Blue and purple bars correspond to primary and secondary analyses, respectively. Error bars indicate standard error of the estimates.

Table 1: Sample characteristics and morphometricity estimates for analyzed disease traits.

*P < 5e-3, **P < 5e-4, ***P < 5e-5

Disease	Case N	Case Age (mean±std)	Case Fem%	Control N	Control Age (mean±std)	Control Fem %	Study Name	Assumed Prevalance	Morphometricity (liability scale)	Standard Error
Alzheimer's	154	74.6±7.6	46.8	219	75.9±5	48.4	ADNI	13%	1.00***	0.03
ADHD	122	11.6±3.2	75.4	384	11.8±2.9	62.2	ADHD	1%	0.55**	0.16
Schizophrenia	92	33±11.2	75.0	85	32.8±11.7	67.1	MCIC	1%	0.50***	0.03
Autism	209	17.6±7.9	14.4	305	17.3±7.2	18.4	ABIDE	1%	0.38***	0.06
Parkinson's	376	61.4±9.7	35.1	152	60.4±11.4	47.8	PPMI	0.2%	0.20*	0.06

Table 2: Sample characteristics and morphometricity estimates for the analyzed general nonclinical traits.

Trait Name	Sample Size	Age (mean±std) [min-max]	Female %	Study Name	Morhopmetricity Estimate	Standard Error
Age	1073	22±5.8 [18-81]	56	GSP	1.00	0.01
IQ	155	11.3±2.8 [7.2-17.7]	60	ADHD	0.95	0.05
Sex	1074	25.3±13.7 [18-84]	50	GSP	0.93	0.02
Education	152	60.4±11.4 [30-82]	34.8	PPMI	0.81	0.08

References

1. Smith, C. D., Chebrolu, H., Wekstein, D. R., Schmitt, F. A., & Markesbery, W. R. (2007). Age and gender effects on human brain anatomy: a voxel-based morphometric study in healthy elderly. *Neurobiology of aging*, *28*(7), 1075-1087.
2. Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Büchel, C., & May, A. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. *The Journal of Neuroscience*, *26*(23), 6314-6317.
3. Thompson, P. M., Cannon, T. D., Narr, K. L., Van Erp, T., Poutanen, V. P., Huttunen, M., ... & Dail, R. (2001). Genetic influences on brain structure. *Nature neuroscience*, *4*(12), 1253-1258.
4. Thompson, P. M., Ge, T., Glahn, D. C., Jahanshad, N., & Nichols, T. E. (2013). Genetics of the connectome. *Neuroimage*, *80*, 475-488.
5. Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and affective neuroscience*, *2*(1), 67-70.
6. Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). *Statistical parametric mapping: the analysis of functional brain images: the analysis of functional brain images*. Academic press.
7. Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage*, *11*(6), 805-821.
8. Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., ... & Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral cortex*, *14*(7), 721-730.
9. Avants, B.B., Cook, P.A., Ungar, L., Gee, J.C., & Grossman, M. (2010) Dementia induces correlated reductions in white matter integrity and cortical thickness: a multivariate neuroimaging study with sparse canonical correlation analysis. *Neuroimage*, 50(3):1004-16.
10. McIntosh, A.R., Lobaugh, N.J. (2004) Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage*, 23:S250-63.
11. Friston, K., Chu, C., Mourão-Miranda, J., Hulme, O., Rees, G., Penny, W., & Ashburner, J. (2008). Bayesian decoding of brain images. *NeuroImage*, *39*(1), 181-205.
12. Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*, *45*(1), S199-S209.

13. Sabuncu, M. R., Konukoglu, E., & Alzheimer's Disease Neuroimaging Initiative. (2015). Clinical prediction from structural brain MRI scans: a large-scale empirical study. *Neuroinformatics*, *13*(1), 31-46.
14. Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, *87*, 96-110.
15. Wray, N., & Visscher, P. (2008). Estimating trait heritability. *Nature Education*, *1*(1), 29.
16. Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255-266.
17. Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., ... & Goddard, M. E. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature genetics*, *42*(7), 565-569.
18. Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, *88*(1), 76-82.
19. Ge, T., Nichols, T. E., Lee, P. H., Holmes, A. J., Roffman, J. L., Buckner, R. L., ... & Smoller, J. W. (2015). Massively expedited genome-wide heritability analysis (MEGHA). *Proceedings of the National Academy of Sciences*, *112*(8), 2479-2484.
20. Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*(2), 774-781.
21. Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, *58*(3), 545-554.
22. Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, *72*(358), 320-338.
23. Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits* (Vol. 1). Sunderland, MA: Sinauer.
24. Dempster, E. R., & Lerner, I. M. (1950). Heritability of threshold characters. *Genetics*, *35*(2), 212-236.
25. Lee, S. H., Wray, N. R., Goddard, M. E., & Visscher, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *The American Journal of Human Genetics*, *88*(3), 294-305.

26. Holmes, A. J., Hollinshead, M. O., O'Keefe, T. M., Petrov, V. I., Fariello, G. R., Wald, L. L., ... & Smoller, J. W. (2015). Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific data*, *2*.
27. Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn human connectome project: an overview. *Neuroimage*, *80*, 62-79.
28. Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., ... & Dale, A. M. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*, *27*(4), 685-691.
29. Milham, M. P., Fair, D., Mennes, M., & Mostofsky, S. H. (2012). The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in systems neuroscience*, *6*, 62.
30. Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, *19*(9), 1498-1507.
31. COBRE. http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html.
32. Gollub, R. L., Shoemaker, J. M., King, M. D., White, T., Ehrlich, S., Sponheim, S. R., ... & O'Leary, D. (2013). The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics*, *11*(3), 367-388.
33. Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., ... & Deen, B. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, *19*(6), 659-667.
34. Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., ... & Poewe, W. (2011). The Parkinson progression marker initiative (PPMI). *Progress in neurobiology*, *95*(4), 629-635.
35. Alzheimer's Association. (2014). 2014 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *10*(2), e47-e92.
36. Hebert, L. E., Scherr, P. A., Bienias, J. L., Bennett, D. A., & Evans, D. A. (2003). Alzheimer disease in the US population: prevalence estimates using the 2000 census. *Archives of neurology*, *60*(8), 1119-1122.

37. Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and meta-regression analysis. *American journal of psychiatry*.
38. Lee, S. H., DeCandia, T. R., Ripke, S., Yang, J., Sullivan, P. F., Goddard, M. E., ... & International Schizophrenia Consortium. (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature genetics*, *44*(3), 247-250.
39. Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(8), 921-929.
40. Keller, M. F., Saad, M., Bras, J., Bettella, F., Nicolaou, N., Simón-Sánchez, J., ... & Schulte, C. (2012). Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. *Human molecular genetics*, *21*(22), 4996-5009.
41. Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics*, *45*(9), 984-994.
42. Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
43. Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., & Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron*, *62*(1), 42-52.
44. Huttenlocher, P. R. (1994). Synaptogenesis, synapse elimination, and neural plasticity in human cerebral cortex. *Threats to optimal development: Integrating biological, psychological, and social risk factors*, *27*, 35-54.
45. Lewis, D. A., & Levitt, P. (2002). Schizophrenia as a disorder of neurodevelopment. *Annual review of Neuroscience*, *25*(1), 409-432.
46. Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on psychological science*, *4*(3), 274-290.
47. Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature neuroscience*, *12*(5), 535-540.
48. Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B., Sabuncu, M. R., & Alzheimer's Disease Neuroimaging Initiative. (2013). Statistical analysis of longitudinal neuroimage data with linear mixed effects models. *Neuroimage*, *66*, 249-260.

49. Bernal-Rusiel, J. L., Reuter, M., Greve, D. N., Fischl, B., Sabuncu, M. R., & Alzheimer's Disease Neuroimaging Initiative. (2013). Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. *NeuroImage*, *81*, 358-370.
50. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M., & Wray, N. R. (2012). Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*, *28*(19), 2540-2542.
51. Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... & Montillo, A. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341-355.
52. Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... & Albert, M. S. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-980.
53. Liu, D., Lin, X., & Ghosh, D. (2007). Semiparametric Regression of Multidimensional Genetic Pathway Data: Least-Squares Kernel Machines and Linear Mixed Models. *Biometrics*, *63*(4), 1079-1088.
54. Molenberghs, G., & Verbeke, G. (2007). Likelihood ratio, score, and Wald tests in a constrained parameter space. *The American Statistician*, *61*(1), 22-27.
55. Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., ... & Nolan, D. (2013). Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage*, *80*, 169-189.

Supporting Information for "*Morphometricity: Measuring the Neuroanatomical Signature of a Trait*"

Mert R. Sabuncu, Tian Ge, Avram J. Holmes, Jordan W. Smoller, Randy L. Buckner, and Bruce Fischl; for the Alzheimer's Disease Neuroimaging Initiative (ADNI)*

Supplementary Table S1: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values (Equation 6) for the primary morphometricity analyses conducted with the Linear (Equation 4) and Gaussian (Equation 5) kernels.

	AIC ($x10^4$)		**BIC** ($x10^4$)	
Disease/Trait	Gaussian	Linear	Gaussian	Linear
Alzheimer's (ADNI)	-1.508	-0.309	-1.362	-0.163
ADHD	-2.040	-0.655	-1.826	-0.441
Schizophrenia (MCIC)	-0.772	-0.216	-0.716	-0.160
Autism (ABIDE)	-2.022	-0.637	-1.804	-0.419
Parkinson's	-0.522	-0.194	-0.491	-0.163
Age (GSP)	-2.967	-0.276	-2.433	0.258
IQ (ADHD 200)	-0.521	-0.107	-0.474	-0.059
Sex (GSP)	-3.559	-0.992	-3.024	-0.457
Education (PPMI)	-0.512	-0.188	-0.477	-0.153

Supplementary Table S2: Parametric and jackknife estimates of morphometricity and standard errors. Morphometricity estimates for disease traits are on the liability scale.

	Morphometricity		Standard Error	
Disease/Trait	Parametric	Jackknife	Parametric	Jackknife
Alzheimer's (ADNI)	1.00	1.00	0.03	0.03
ADHD (ADHD 200)	0.55	0.55	0.16	0.17
Schizophrenia (MCIC)	0.50	0.50	0.03	0.03
Autism (ABIDE)	0.38	0.38	0.06	0.06
Parkinson's	0.20	0.20	0.06	0.06
Age (GSP)	1.00	1.00	0.01	0.01
IQ (ADHD 200)	0.95	0.95	0.05	0.06
Sex (GSP)	0.93	0.93	0.02	0.02
Education (PPMI)	0.81	0.81	0.08	0.07

Supplementary Table S3: Independent replication on non-overlapping data: sample characteristics and morphometricity estimates for analyzed disease traits.

Disease	Case N	Case Age (mean±std)	Case Fem%	Control N	Control Age (mean±std)	Control Fem %	Study Name	Assumed Prevalance	Morphometricity (liability scale)	Standard Error
Alzheimer's	100	76.8±7.1	59.0	98	75.9±9.0	73.4	OASIS	13%	1.00	0.06
Schizophrenia	56	35.8±12.4	16.1	73	35.7±11.6	31.5	COBRE	1%	0.45	0.08

Supplementary Table S4: Independent replication on non-overlapping data: sample characteristics and morphometricity estimates for analyzed general nonclinical traits.

Trait Name	Sample Size	Age (mean±std) [min-max]	Female %	Study Name	Trait (mean±std) [min-max]	Morhopmetricity Estimate	Standard Error
Age	124	68.0±13.7 [33-94]	72.6	OASIS	N/A	1.00	0.03
IQ	108	17.5±8.0 [10.5-56.2]	14.8	ABIDE	109.33±11.0[84-133]	0.95	0.06
Sex	298	11.3±2.9 [7.17-21.74	50	ADHD	N/A	0.90	0.05
Education	124	68.0±13.7 [33-94]	72.6	OASIS	3.5±1.2 [1-5]	0.79	0.13

Supplementary Table S5: ROI-based morphometricity estimates for the traits with two independent samples. Most associated ROIs were identified on independent sample. *For disease traits we list morphometricity estimates on the liability scale.

Trait Name	ROI-based Morphometricity	Standard Error	P-value	Most associated ROI
Age	0.82	0.03	<1e-15	3rd Ventricle
IQ	0.87	0.05	<1e-15	Right Lateral Ventricle
Sex	0.44	0.07	1.9e-9	Right Thalamus
Education	0.00	0.06	1.00	Right Hippocampus
Alzheimer's*	0.66	0.16	2.6e-5	Right Hippocampus
Schizophrenia*	0.06	0.05	0.22	Right Inferior Lateral Ventricle