On Statistical Analysis of Brain Variability*

In memory of Yao Suzhen (1929 - 2021)

Oliver Y. Chén[1,2†], Huy Phan[3], Guy Nagels[4], and Maarten de Vos[5]

[1]Faculty of Social Sciences and Law, University of Bristol, Bristol, UK.

[2]Division of Biosciences, University College London, London, UK.

[3]Department of Computer Science, Queen Mary University of London, London, UK.

[4]Department of Neurology, Universitair Ziekenhuis Brussel, Brussel, Belgium.

[5]Faculties of Medicine and Engineering Science, KU Leuven, Leuven, Belgium.

*An early version of this paper was posted in *preprints.org* in 2020 (doi: 10.20944/preprints202008.0428.v1).

[†]Correspondence to: olivery.chen@bristol.ac.uk.

Abstract

We discuss what we believe could be an improvement in future discussions of the ever-changing brain. We do so by distinguishing different types of brain variability and outlining methods suitable to analyse them. We argue that, when studying brain and behaviour data, classical methods such as regression analysis and more advanced approaches both aim to decompose the total variance into sensible variance components. In parallel, we argue that a distinction needs to be made between innate and acquired brain variability. For varying high-dimensional brain data, we present methods useful to extract their low-dimensional representations. Finally, to trace potential causes and predict plausible consequences of brain variability, we discuss how to combine statistical principles and neurobiological insights to make associative, explanatory, predictive, and causal enquires; but cautions are needed to raise association- or prediction-based neurobiological findings to causal claims.

Prologue

Darwin discussed the importance of variability in *On the Origin of Species* and argued it is greatest in structures that evolve fastest [1]. In humans, the brain is the most variable organ [2]. As a knowledge-acquiring system, the human brain seeks to extract invariant, permanent, and unchanging information from its environment, which enables learning and surviving [3–5]. The inspection of cytoarchitecture by Campbell and Brodmann unveiled the brain's varying organization and functioning [6–8]. The introduction of *variance* by R.A. Fisher launched the quantitative enquiry of biological variability [9].

During the past century, linking covarying neural features and behavioural measurements, scholars have uncovered remarkable insights about the brain, mind, and behaviour [10,11,20,12–19]. In parallel, the analysis of variability using neural, cognitive, and disease data has provided plausible explanations on the neural origins of cognition and behaviour [21–23] and discovered potential markers predictive of brain disorders [24,25].

The major theme of this paper is to connect two areas of equal importance: on the one hand the biological variability of the brain and on the other hand the statistical methods and applications useful to study it. We hope our explorations may stir further discussions about neurobiological underpinnings of brain variability and reliable and reproducible methods to study it.

As a preamble, we outline the topics covered in this paper:

i. We define different types of brain variability.
ii. We argue that, when studying brain and behaviour data, classical statistical methods and more advanced approaches aim at decomposing the total variance into sensible variance components.
iii. We argue that a distinction needs to be made between innate and acquired brain variability.
iv. We suggest methods to obtain low-dimensional representations from varying high-dimensional brain data.
v. We discuss associative, explanatory, predictive, and causal analyses of the varying brain.

1. Defining brain variability

To define different types of variability, let's consider N subjects each of whose brain consists of $\boldsymbol{V}$ areas measured along T timepoints. Let $y_i(\boldsymbol{v}, t)$ denote the signal measured at area $\boldsymbol{v} \in \{\boldsymbol{1}, \ldots, \boldsymbol{V}\}$ at time $1 \leq t \leq T$ from an individual $1 \leq i \leq N$. We consider the location and types of signals broadly: the former can be a neuron, a voxel, or a brain parcel; the latter can be the action potentials of single neurons, BOLD fMRI of voxels, or EEG recordings of electrodes.

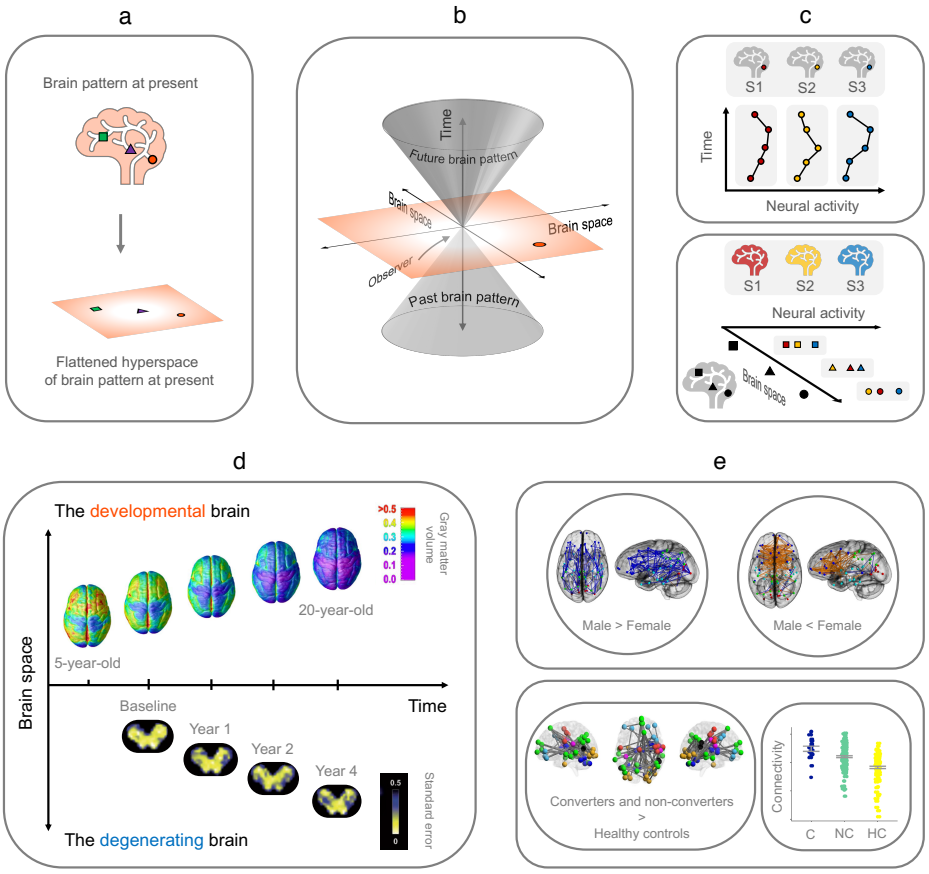

Figure 1. Distinguishing types of brain variability. (a) A brain fixed in time. 3-D brain data observed at each time point are flattened into a 2-D matrix. **(b) A brain travelling in time.** Suppose an observer is at the centre of the 2-D matrix at $t = 0$. Looking around, the observer sees neural activities across the entire brain at present. Looking below and above, the observer sees past and future brain activities. **(c) Temporal, spatial, and within-group brain variability.** Top: The variability for a fixed brain area recorded over time shows temporal variation. Bottom: The variability across different brain areas depicts spatial variation. If we fix both space and time, the variability among individuals in the group forms within-group variation. **(d) The developmental and the degenerating brain.** One can study how the brain develops and degenerates by tracing brain activities over time during childhood and adolescence [26] and by following patients with neurodegenerative diseases [27]. **(e) The between-group brain variability.** Variability between different gender [28] and disease [24] groups present between-group variation.

A powerful way to think of the varying brain is to see it as a three-dimensional brain data cubic travelling in time (**Figure** 1 a-b). Following signals from subject i's

brain area $\boldsymbol{v}$ over time, the trajectory $\left(y_i(\boldsymbol{v},1), y_i(\boldsymbol{v},2), \ldots, y_i(\boldsymbol{v},T)\right)$ shows *temporal variation*[i] (top panel of **Figure** 1 c). If we fix time t, the distribution of signals across different brain areas $\left(y_i(\mathbf{1},t), y_i(\mathbf{2},t), \ldots, y_i(\boldsymbol{V},t)\right)$ depicts *spatial variation* (bottom panel of **Figure** 1 c). If we fix both space $\boldsymbol{v}$ and time t, the distribution $\left(y_1(\boldsymbol{v},t), y_2(\boldsymbol{v},t), \ldots, y_N(\boldsymbol{v},t)\right)$ presents *within-group variation* among N individuals (**Figure** 1 c). Finally, comparing patterns between groups, for example, male $\left(y_1^M(\boldsymbol{v},t), y_2^M(\boldsymbol{v},t), \ldots, y_{N_1}^M(\boldsymbol{v},t)\right)$ *vs.* $\left(y_1^F(\boldsymbol{v},t), y_2^F(\boldsymbol{v},t), \ldots, y_{N_2}^F(\boldsymbol{v},t)\right)$ female (or healthy *vs.* disease), we witness *between-group variation* (**Figure** 1 e).

The differentiation of brain variability drives us into specialized areas of brain study (**Figure** 1). **Temporal variability**: Tracking temporal brain variability, one gains insights into time-varying neural dynamics (*e.g.*, "dynamic core" [29]), neural development, and brain maturation [4,26,30]. Analysing temporal brain variability for elderlies or patients with neurodegenerative diseases helps to understand the ageing brain [31,32] and the (neuro)degenerating brain [27]. Additionally, past neural activities help to make forecasts about future activities [33]. **Spatial variability**: Studying brain areas whose signals co-vary in space helps to decipher how the brain is wired [34,35]. **Within- and between-group variability**: Examining brain patterns within and across groups of individuals, one can derive population-level characteristics [28,36] and subject-specific information for group- or patient-identification [24]. **Potential causal variability**: Examining brain areas whose patterns co-vary with external stimuli and/or behaviour helps to identify neural signatures: specialized for processing the stimuli [37], predictive of behaviour [24,38–41], and intermediating stimuli and behaviour [42,43].

2. Identifying, isolating, and quantifying brain variability

Generally, brain variability consists of temporal, spatial, individual-, and population-level sources, and may also be affected by covariates such as age and gender. Here, we present a simple way to identify, isolate, and quantify sensible variance components via variance decomposition.

2.1 Variance decomposition via ANOVA and ANCOVA

[i] The temporal variance is: $\sum_{t=1}^{T}\left(y_i(\boldsymbol{v},t) - \bar{y}_i(\boldsymbol{v})\right)^2$, where $\bar{y}_i(\boldsymbol{v}) = \frac{\sum_{t=1}^{T} y_i(\boldsymbol{v},t)}{T}$. We omit similar calculations henceforth.

Consider a bi-variate ANOVA model [ii] : $y_i(\boldsymbol{v},t) = \mu + \alpha_i + s(\boldsymbol{v}) + \tau(t) + (s\tau)_{\boldsymbol{v}t} + \varepsilon_i(\boldsymbol{v},t)$, where μ indicates the group-level mean, α_i denotes the subject-specific departure from the mean, $s(\boldsymbol{v})$ stands for the spatial deviates of area $\boldsymbol{v}$, $\tau(t)$ represents the longitudinal fluctuation, $(s\tau)_{\boldsymbol{v}t}$ is the space-time interaction, and $\varepsilon_i(\boldsymbol{v},t)$ designates the residual where $\varepsilon_i(\boldsymbol{v},t) \sim N(0,\sigma^2)$ (**Figure** 2 and **Table** 1 in **Supplementary Materials**). The total sum of squared residuals (SST) can be decomposed[iii] into: (a) an individual component (SS_α), (b) a brain space component (SS_{Space}), (c) a temporal component (SS_{Time}), and (d) an error component (SSE).

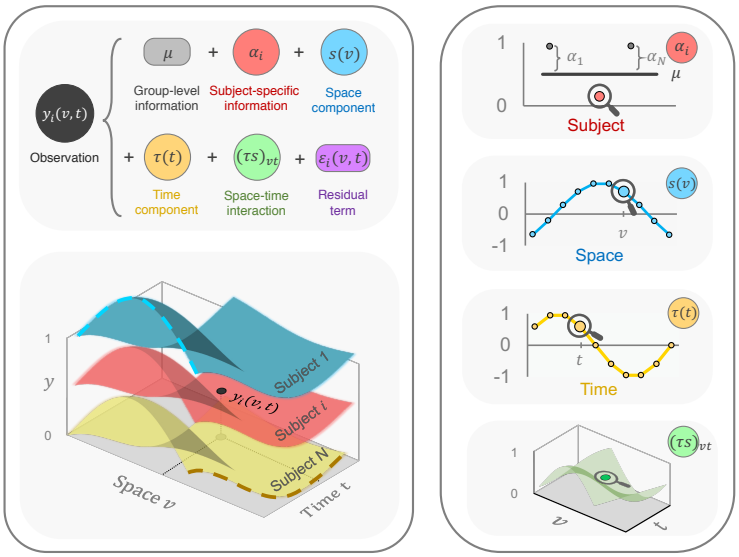

Figure 2. A variance-decomposition view of regression analysis. Top left: A bi-variate function consists of a group-level mean, a subject-specific deviation, a temporal fluctuation, a spatial shift, a spatial-temporal interaction, and the residual term. **Bottom left:** Surfaces of three individuals. One point $y_i(v,t)$ is highlighted for individual i at location v at time t. **Right:** Sources of variation. From the top to bottom are: (1) group-level mean μ and a subject-specific deviation α_i, (2) spatial shift $s(v)$, (3) temporal fluctuation $\tau(t)$, and (4) spatial-temporal interaction $(\tau s)_{vt}$.

2.2 Further variance decomposition on the residuals

The residual may contain a portion of total variance for which has not been accounted by existing model settings. One way to examine the remaining variability is to further

[ii] One can extend it to an ANCOVA: $y_i(v,t) = \mu + \alpha_i + s(v) + \tau(t) + (s\tau)_{vt} + \boldsymbol{z}_i\boldsymbol{\beta} + \varepsilon_i(v,t)$, where $\boldsymbol{z}_i$ denotes covariates and $\boldsymbol{\beta}$ represents their parameters.

[iii] Specifically, $SST = \sum_{i=1}^{N}\sum_{v=1}^{V}\sum_{t=1}^{T}(y_i(v,t) - \bar{y}_{...})^2$, $SS_\alpha = VT\sum_{i=1}^{N}(\bar{y}_{i..} - \bar{y}_{...})^2$ with $N-1$ degrees of freedom (df), $SS_{Space} = NT\sum_{v=1}^{V}(\bar{y}_{.v.} - \bar{y}_{...})^2$with $V-1$ df, $SS_{Time} = NV\sum_{t=1}^{T}(\bar{y}_{..t} - \bar{y}_{...})^2$ with $T-1$ df, $SS_{Space/Time} = N\sum_{t=1}^{T}\sum_{v=1}^{V}(\bar{y}_{.vt} - \bar{y}_{...})^2$ with $(N-1)(V-1)$ df, and $SSE = \sum_{i=1}^{N}\sum_{v=1}^{V}\sum_{t=1}^{T}\left(y_i(v,t) - \hat{y}_i(v,t)\right)^2$, with $[NVT - (N-1)(V-1) - (N-1) - (V-1) - (T-1) - 1]$ df, where $\hat{y}_i(v,t) = \hat{\mu} + \hat{\alpha}_i + \hat{s}(v) + \hat{\tau}(t) + (\hat{s\tau})_{vt}$.

decompose it. To demonstrate this, let's extend the model in **Section** 2.1 to a general form[iv]:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon}$ is the residual following $MVN(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} = \text{blockdiag}(\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n)$, and $\boldsymbol{V}_i$ models the within-subject correlation structure, for $1 \leq i \leq N$. The model is general because if the outcome variable is the brain data (as in **Section** 2.1), the design matrix $\boldsymbol{X}$ includes the individual, temporal, spatial, interaction, and covariate entries; if the outcome variable is measured behaviour, and one wants to study how brain signals affect behaviour controlling for other effects, the design matrix includes brain signals in addition to other parameters.

To avoid distractions (*e.g.*, specifying all choices of the design matrix and parameter estimation), let's directly look at the residual $\boldsymbol{\epsilon}$. It can be decomposed into three parts: random effect, serial correlation, and measurement error [44] (**Table** 2 in **Supplementary Materials**). One can interpret the residual variability using newly separated components[v].

3. Distinguishing innate and acquired brain variability

As a biological organ, the brain's structure and functioning are, in part, endorsed by innate factors. Living in an ever-changing environment, the brain adapts to the external world [3–5]. But how do we distinguish innate and acquired brain variability? We draw insights from three directions.

From biology to neurobiology. There have been extensive discussions about innate *vs.* acquired biological entities [45–51]. Generally, the former include genetic and epigenetic information [52]. The latter include environmental factors [53], nutrition and diet [54], and disease [55].

A direct message from general biology is that there is a need to distinguish innate and acquired brain variability [5] (**Figure** 3). The innate brain variability is likely dictated by the genes and is less variable (compared to the acquired variability). The

[iv] To recapitulate a non-linear relationship, the model can be modified to a generalized estimation equation.

[v] The decomposed residual term may not contain all three parts. Models whose residual term only contains measurement error are called marginal models; models whose error term contains random effect and measurement error are called random effect models; models whose error term contains serial effect and measurement error are called transition (or Markov) models. The error may still contain a smaller amount of unexplained variability.

acquired brain variability is developed postnatally, due to environmental factors or a combination of environmental and genetic factors and is more variable. Next is to find evidence. Colour perception varies little in humans. Perceiving white colour after seeing a white flag is independent of culture and learning [56,57]. Perceiving ceasefire or surrender when seeing a white flag, however, depends on postnatal learning; it is more variable across different ages and cultural groups [5]. Similarly, face recognition is potentially innate, while associating faces of different races with social categorization stereotyping, and discrimination is perhaps acquired [58].

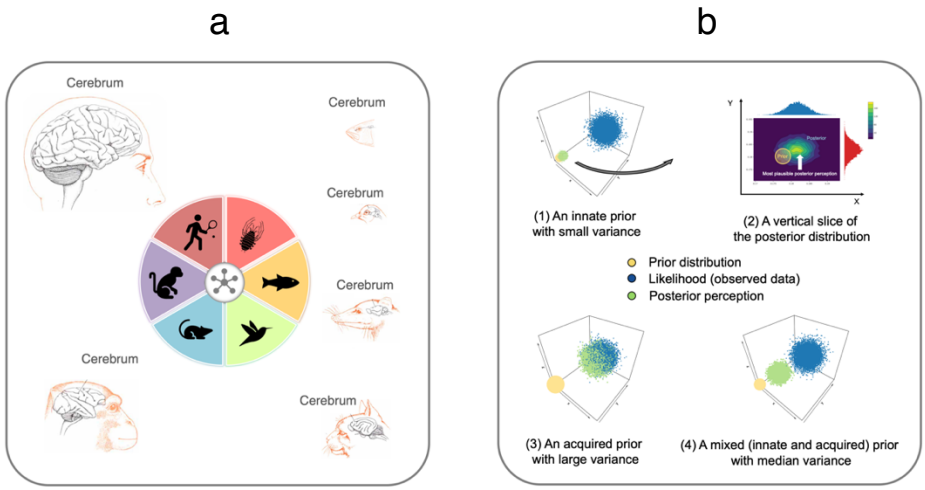

Figure 3. Two views of brain variability. (a) A neurobiological view. Empirical evidence for Bayesian updating has been found in amphipods, fish, birds, small mammals, monkeys, and humans (see text for details). The cerebra of animals and humans are shown to indicate brain sizes. Cerebra images are redrawn from "The Brain" by David H. Hubel. Copyright Patricia J. Wynne 1979. **(b) A data science view.** Priors, observed data, and posterior are represented in yellow, blue, and green, respectively. The size of circles indicates the extent of variability. (1) When the prior distribution has very small variability (*i.e.*, high precision), the posterior distribution is derived mainly from the prior, and not much from the observed data. This constitutes the concept of an innate prior. (2) shows that, with innate priors, the centre of the posterior distribution is very close to that of the prior distribution. (3) shows that when the prior distribution has very large variability, the posterior distribution is modified mainly by the observed data. This constitutes the concept of an acquired prior. (4) shows that when the prior distribution has moderate variability, the posterior distribution is derived from both the prior and the observed data; the centre of the posterior distribution is between the prior distribution and the distribution of the data. Image is adapted by permission from *The European Journal of Neuroscience* "The Bayesian-Laplacian Brain" by S. Zeki and O.Y. Chén. Copyright 2019.

From statistics to neurobiology. The variance decomposition (**Section** 2) enables one to separate the total variability of the (neural and behavioural) phenotypes into components corresponding to genetic, environmental, and interaction factors. The interaction can be categorized into gene-gene or environment-environment interaction [59] and gene-environment interaction [60–62].

A derivative of variance decomposition is the heritability analysis. Let $P = G + E$ denote the relationship between a phenotype (P) and its genetic (G) and

environmental (E) factors. The heritability (H^2) is the fraction of genetic variance over the total variance, namely $H^2 = \frac{V(G)}{V(P)}$, where the total variance is the sum of genetic variance, variance due to environmental factors, and twice the covariance between genetic and environmental factors.

From innate and acquired brain variability to the Bayesian brain. Empirical studies suggest that small animal brains perform Bayesian updating, integrating prior information and postnatal learning (**Figure** 3) [63–66]. Similar behaviour is also seen in small mammals [67]. Monkeys estimate time by integrating sensory evidence with prior beliefs [68]. Humans use probability updating to modify their perception [69], cognition [70], and sensorimotor function [71]. Incorporating prior knowledge (encoded in genes and acquired postnatally through learning) with new information, perception and behaviour are updated with higher precision (*i.e.*, lower variation) [5].

4. Analysis of variability using big brain data

How can we extract generalisable knowledge and principles from the varying large-scale, potentially high-dimensional brain data? Here, we first present the Neural Law of Large Numbers (NLLN) that may offer insights into deriving population patterns or general principles. Next, we review methods suitable for extracting low-dimensional representations from varying high-dimensional brain data. Finally, we discuss challenges studying big brain data.

4.1 The neural law of large numbers (NLLN)

The NLLN states that the average patterns of varying brain signals may provide insights about population traits or general neurobiological principles.

Single neurons. When a black bar moved at different orientations in front of a macaque monkey, a single neuron from the striate cortex fired. Although the neuron responded to motions from several directions (**Figure** 4 a), it responded, on average, most strongly to the up-right motion, suggesting it was motion-selective [72].

When white and monochromatic light of different wavelengths was present to a macaque monkey, a single neuron from layer 2 of the striate cortex fired. Under the same motion, this neuron responded to colour of different wavelengths; but it responded, on average, most excitingly to a wavelength of 450mμ, suggesting it was colour-selective (**Figure** 4 b) [72].

Populations of neurons. When a monkey viewed faces and non-face objects, neurons in the inferior temporal cortex showed responses to all stimuli. But the average responses were the largest when the monkey viewed normal monkey and human faces, suggesting these neurons were involved in face recognition (**Figure** 4 c) [73].

Brain networks. Brain connectivities follow the NLLN. This is observed from connectivity between a pair of brain areas, connectivities between one brain area and the rest of the brain, and across the whole-brain (**Figure** 4 d-f).

The NLLN does not imply the asymptotic mean equals to the population truth; rather, it suggests that averaged features may provide relatively stable and reliable estimates for population patterns and general principles of the varying brain data, subject to noises and errors[vi].

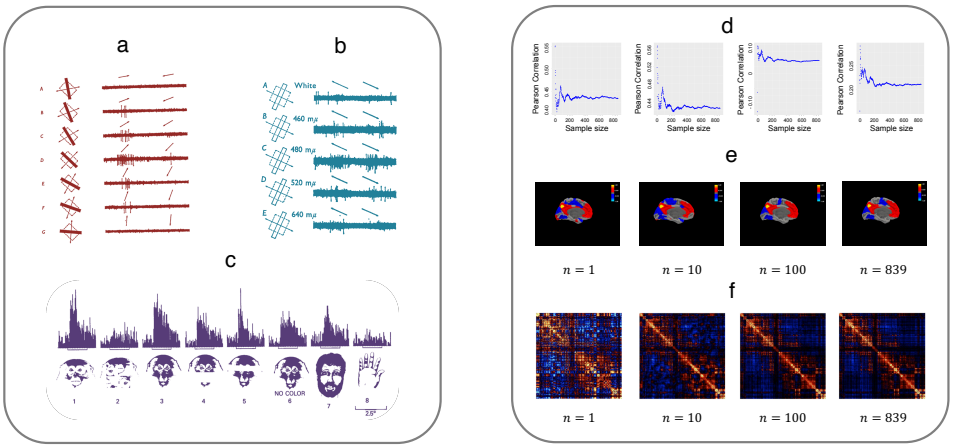

Figure 4. The Neural Law of Large Numbers (NLLN). (a) A motion-selective neuron. When a black bar moves at different orientations in front of a macaque monkey, a single neuron from the striate cortex is excited. This neuron responds, on average, best to up-right motion. **(b) A colour-selective neuron.** When white and monochromatic light of different wavelengths is present to a macaque monkey, a single neuron from layer 2 of the striate cortex is excited. Under the same motion, this neuron responds, on average, the best to a wavelength of $450m\mu$. **(c) Neurons for face recognition.** When a monkey viewed a monkey face, a human face, fractions of a monkey face, an abnormal monkey face, and a non-face object, neurons in the inferior temporal cortex respond, on average, the best to normal monkey and human faces. **(d)-(f) Brain connectivity.** From top to bottom are connectivities between paired brain areas, connectivities between a seed (in the posterior cingulate; yellow coulour) and the rest of the brain, and the whole-brain connectivity matrices. Images in (a) and (b) are reproduced by permission from RightsLink *Journal of Physiology* "Receptive fields and functional architecture of monkey striate cortex" by T. N. Wiesel and D. H. Hubel (Copyright 1968). Image in (c) is reproduced from "Stimulus-selective properties of inferior temporal neurons in the macaque" by Desimone *et al.* (Copyright 1984 *Society for Neuroscience*). Figures (d)-(f) are produced using data from the Human Connectome Project [74].

[vi] By taking the average, the random noises are typically zero centred, but there are still systematic biases and measurement errors.

4.2 High-dimensional brain data analysis

A type of big brain data requiring special treatments is the high-dimensional brain data, particularly when the number of features (p) goes to infinity faster than (or at the same rate as) the number of sample size (N). Under these circumstances, the classical statistical theories collapse[vii].

The difficulty, however, can be partially alleviated if there exists a sparse structure. Under this assumption, one can deploy feature selection or transformation to discover a sparse or low-dimensional representation. This assumption is not unrealistic: empirical studies have unveiled sparse structures in cells [75], genes [76], and the brain [77].

There are, generally, two types of feature selection: stepwise feature selection and regularization [78]. When p grows faster or as fast as N, regularized models may not pick up the correct features. The sure independence screening addresses this issue by reducing the dimensionality to a moderate number $d < N$ before applying regularization [79]. When the relationship between brain features and the outcome is nonlinear, one can consider sparse additive models [80] and Bayesian additive regression trees [81].

To extract low-dimensional representations from high-dimensional data, one can perform feature transformation such as eigenanalysis[viii] or look for a low-dimensional nonlinear manifold[ix] [82].

4.3 Further challenges and potential solutions

Big brain data may confront **spurious correlations**[x]. Consider *p* neural edges. If *p* is large, it is likely that a few edges are *spuriously*[xi] associated with an outcome in a sample. Out-of-sample prediction, cross-validation, and repeated sampling test may alleviate this issue.

[vii] Classical large sample theories focus on scenarios where p is smaller than n.

[viii] Including principal components analysis, canonical correlation analysis, multivariate analysis of variance, and discriminant analysis. Direct thresholding [110] and sparse covariance models [111] are useful for exploring the sparsity nature of a large covariance matrix.

[ix] Including ISOMAP [112], locally linear embedding [113], Laplacian eigenmaps [114], and Hessian eigenmaps [115].

[x] Like spurious correlation, endogeneity can also be caused by (uncontrolled) confounding variables, or by simultaneity (*i.e.*, a looped causal effect between the features and the outcome).

[xi] Another source of spurious correlation is due to confounding variables.

Related to spurious correlation is **incidental endogeneity**, where some features are coincidentally correlated with the residual term. Most statistical models require the predictors are uncorrelated with the residual, which may be violated considering high-dimensional brain data. The treatment for incidental endogeneity is actively pursued (see [83] for a solution under assumptions).

Next, aggregating datasets of various levels of **noises** may bias the estimates. To address this, one needs rigorous pre-processing and suitable aggregation methods (*e.g.*, weighting the dataset by its inverse of variance).

High-dimensional data face **computational challenges**. Many high-dimensional problems are intractable. Additionally, they may generate even larger intermediate data. For example, some models require inverting a correlation matrix [84] but it may be challenging to invert a correlation matrix calculated from one-million features (but see sparse assumption in **Section** 4.2).

Data with large sample sizes may yield small, yet **significant**, effect sizes. *P-values* in these cases may offer little inference value [85]; the effects may not be meaningful in clinical trials or pathological studies, and are difficult to interpret and reproduce [86].

Finally, data **visualization** is critical to exploratory analysis and *post hoc* interpretation but plotting high-dimensional data is difficult. One can instead project[xii] high-dimensional data onto low-dimensional space (**Figure** S1 in **Supplementary Materials**) [87].

5 Associative, explanatory, predictive, and causal analyses of brain variability

We have so far largely discussed the definition, quantification, and exploratory analysis of brain variability. But what are the causes of brain variability? How does one brain area's variability affect another's? What are the consequences of brain variability? And how do we study the causes and consequences?

Answering this set of questions requires a causal enquiry of the brain. In **Section** 5.1, we discuss how to find potential causal markers from effective connectivity via predictive modelling. In **Section** 5.2, we compare associative,

[xii] Including *principal component analysis*, *random projection* [116], and *t-distributed stochastic neighbour embedding* (t-SNE) [117,118].

explanatory, predictive, and causal analyses of brain variability and propose strategies that may raise associative or predictive findings to (promising) causal discoveries.

5.1 Effective connectivity and predictive modelling

The differentially distributed spatiotemporal variability provides a neural foundation for estimating – via effective connectivity – how signals of one brain area may affect those of others [88,89]. These effective connectivities help to draw a brain atlas consisting of directed neural edges. One can then link those directed edges with measured brain outcomes to find potential neural markers.

Yet it remains possible that some of the effective connectivities are numerical coincidences. A useful way to guard spurious causal claims on effective connectivity is to perform **causal alternation** via, for example, the transcranial magnetic stimulation (TMS), to alter brain signals and check whether the established causal relationship changes (**Section** 5.2).

When causal alternation is inaccessible, one may perform **predictive modeling** to verify if the discovered relationships can be reproduced and extrapolated. Although predictive modeling has made remarkable strides in identifying potential brain areas and pathways linked to perception, motion, and cognition, it does not endorse causation. We discuss this below.

5.2 Comparing associative, explanatory, predictive, and causal analysis of brain variability

Let's begin by defining the problem of causal enquiry [90] regarding the varying brain. Let $\boldsymbol{\xi}$ and ω denote brain features and a univariate brain outcome, respectively. If $\boldsymbol{\xi}$ causally affects ω, we say $\omega = \phi(\boldsymbol{\xi})$, where ϕ denotes a causal map.

As $\boldsymbol{\xi}$ and ω are random variables, one can only aim at uncovering the underlying relationship by analysing their observations, denoted as $\boldsymbol{X}$ and o, respectively. The model is thus $o = f(\boldsymbol{X}) + \varepsilon$, where ε denotes the residual term following a specific distribution.

Using these terminologies, below we compare associative, predictive, explanatory, and causal analyses of brain variability. We omit the comparison between association and prediction, and it between association and explanation, as they are relatively straightforward.

Association *vs.* causation. The estimated relationship $\hat{f}$ suggests only an association between $\boldsymbol{X}$ and o. The reason is twofold. First, due to noise and measurement error, $\boldsymbol{X}$ and o may not accurately describe $\boldsymbol{\xi}$ and ω. Second, even if $\boldsymbol{X}$ and o accurately describe $\boldsymbol{\xi}$ and ω, since f is arbitrary, there may exist a "better" model g ($o = g(\boldsymbol{X}) + e$) that produces a smaller error e than ε (in $o = f(\boldsymbol{X}) + \varepsilon$) or yields better out-of-sample predictions.

Generally, it is reasonable to claim association (at most potential causation) during modeling of f and look for further (*e.g.*, neurobiologicals or medical) confirmation. Additionally, one may consider stringent procedures to make better causal claims (**Figure** S2 in **Supplementary Materials**).

Prediction *vs.* Causation. The battle between prediction and causation can perhaps be seen through the bias-variance trade-off. Consider the expected predictive square error is [91]:

$$E\left(o - \hat{f}(\boldsymbol{x})\right)^2$$

$$= \underbrace{E\left(o - f(\boldsymbol{x})\right)^2}_{\text{True variance}} + \underbrace{\left(E\left(\hat{f}(\boldsymbol{x})\right) - f(\boldsymbol{x})\right)^2}_{\text{Bias}} + \underbrace{E\left(\hat{f}(\boldsymbol{x}) - E\left(\hat{f}(\boldsymbol{x})\right)\right)^2}_{\text{Estimation variance}}.$$

Assuming f approximates the unobserved causal map ϕ, the causal enquiry aims at minimizing the bias between the estimated model $\hat{f}$ and f. Predictive modeling aims at minimizing both bias and the estimation variance, even if at a cost of theoretical accuracy for improved empirical precision; in other words, one prefers a "wrong" (or less realistic) model that yields better predictions [90].

Predictive modeling generally reduces overfitting. It *may* (our emphasis) help to raise association to potential causation: if trained $\hat{f}$ can predict outcomes in unseen samples, this suggests $\hat{f}$ represents some general properties preserved in a broad population, hinting a stronger sense of causation (subject to cautions above). Leveraging $\hat{f}$, one can look for a subset of potential causal features. A causal feature, such as a gene mutation [92], is predictive of the outcome and can explain a portion of the total variance. The reverse is not always true.

Explanation *vs.* causation. Until a definitive causal link is charted, one can use predictive or associative analyses to explain the relationship between features and outcome. For example, dexterity features predict Parkinson's disease (PD) in novel samples [93]. But irregular dexterity features are not the cause for PD; they are the consequences of PD. Nevertheless, their predictability is useful for patient

identification and disease severity estimation; they offer insights into the behavioral characteristics of the disease. Additionally, observing patients with motor issues may assist finding neural causes of the PD, for the disruption is likely linked to abnormalities in the motor cortex.

Prediction *vs.* explanation. Predictive models are sometimes built at the cost of explainability. For example, regularization methods reduce estimation variance but introduce bias, making the model less explainable. Ensemble methods improve overall predictability by averaging predictions from individual models, meanwhile the ensembles become difficult to explain [90]. Neural networks may uncover hidden associations between features and outcome and yield accurate predictions, but most are as-of-yet difficult to explain.

Including multiple highly correlated features, such as motor features in PD, or including insignificant (in terms of predictability) features, such as smoking status, may reduce prediction performance, but these features may help to explain the overall problem.

Making better causal inference. Statistical methods useful for evaluating if association can be raised to (promising) causation can be classified into five categories (see [85] for a discussion): (1) randomization [94,95] (and Propensity Score Matching (PSM) [96–99] when randomization is impossible), (2) discovery validation and reproduction [100,101], (3) causal reasoning [102] and, relatedly, graphic models [103–105], (4) causal alternation [106,107], and (5) instrumental variables (IV) [108,109].

Epilogue

Our discussions about brain variability consist of two views: a biological one and a statistical one. Biologically, we discuss the importance of distinguishing different types of brain variability. We suggest a distinction needs to be made between innate and acquired brain variability. Statistically, we argue that regression type of analysis and advanced statistical models aim to decompose the total variance into sensible components attributed to internal or external factors. We review statistical methods to

analyse big brain data and to extract low-dimensional representations from varying high-dimensional brain data.

Uniting biology and statistics, we discuss associative, explanatory, predictive, and causal analyses of the varying brain and suggest strategies that may help to raise association- or prediction-based findings to (promising) causal discoveries.

To conclude, a century ago, Fisher demonstrated that variation among phenotypic traits could be due to Mendelian inheritance [9]. During the past century, the study of variability has time and time again injected fresh insights into brain science and statistical science. Further studies of variability will continue to expand our knowledge about the genetic, environmental, and neural bases of the brain's varying structure and functioning, and how the ever-changing brain makes the ever-evolving humans. ■

Data accessibility. This article has no additional data.

Author's contributions. OYC wrote the manuscript with comments from all other authors.

Competing interests. None declared.

Funding. This research received no external funding.

Acknowledgements. The authors thank Paul Matthews and Tianchen Qian for helpful discussions.

References

1. Darwin C. 1859 On the origin of Species. London, UK: John Murray.
2. Zeki S. 2001 Artistic creativity and the brain. *Science* **293**, 51–52.
3. Edelman GM. 1993 Neural Darwinism: Selection and reentrant signaling in higher brain function. *Neuron* **10**, 115–125.
4. Edelman GM. 1989 *Neural Darwinism: The theory of neuronal group selection*. Oxford, UK: Oxford University Press.
5. Zeki S, Chén OY. 2019 The Bayesian-Laplacian brain. *Eur. J. Neurosci.* **51**,1441–62.
6. Campbell AW. 1905 *Histological studies on the localisation of cerebral function*. *Journal of Mental Science* **50**, 651–662.
7. Brodmann K. 1909 *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig, Germany: Barth.
8. Toro R, Burnod Y. 2005 A morphogenetic model for the development of cortical convolutions. *Cereb. Cortex* **15**, 1900–1913.
9. Fisher RA. 1918 The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans. R. Soc. Edinburgh* **52**, 399–433.
10. Broca P. 1861 Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bull. la Société Anat. Paris* **6**, 330–357.
11. Fritsch G, Hitzig E. 1870 Über die elektrische Erregbarkeit des Großhirns. *Physiol. und Wissenschaftliche Med.* **37**, 300–332.
12. Wernicke C. 1874 *Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis.* Breslau, Poland: Cohn & Weigert.
13. Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS. 1991 A

direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* **11**, 641–649.

14. Zeki S. 1993 *A Vision of the Brain*. Oxford, UK: Blackwell Scientific.
15. Friston K. 2012 The history of the future of the Bayesian brain. *Neuroimage* **62**, 1230–1233.
16. Gordon EM, Laumann TO, Adeyemo B, Petersen SE. 2017 Individual Variability of the System-Level Organization of the Human Brain. *Cereb. Cortex* **27**, 386–399.
17. Halliday DWR, Mulligan BP, Garrett DD, Schmidt S, Hundza SR, Garcia-Barrera MA, Stawski RS, MacDonald SWS. 2017 Mean and variability in functional brain activations differentially predict executive function in older adults: an investigation employing functional near-infrared spectroscopy. *Neurophotonics* **5**, p.011013.
18. Croxson PL, Forkel SJ, Cerliani L, Thiebaut De Schotten M. 2018 Structural Variability Across the Primate Brain: A Cross-Species Comparison. *Cereb. Cortex* **28**, 3829–3841.
19. Seghier ML, Price CJ. 2018 Interpreting and Utilising Intersubject Variability in Brain Function. *Trends Cogn. Sci.* **22**, 517–530.
20. Smith S *et al.* 2019 Structural Variability in the Human Brain Reflects Fine-Grained Functional Architecture at the Population Level. *J. Neurosci.* **39**,6136–6149.
21. Goldman-Rakic PS. 1988 Topography of Cognition: Parallel Distributed Networks in Primate Association Cortex. *Annu. Rev. Neurosci.* **11**, 137–156.
22. Corbetta M, Shulman GL. 2002 Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215.
23. De Felice S, Holland CA. 2018 Intra-individual variability across fluid cognition can reveal qualitatively different cognitive styles of the aging brain. *Front. Psychol.* **9**, 1973.
24. Cao H *et al.* 2018 Cerebello-thalamo-cortical hyperconnectivity as a state-independent functional neural signature for psychosis prediction and characterization. *Nat. Commun.* **9**, 1–9.
25. Christ BU, Combrinck MI, Thomas KGF. 2018 Both reaction time and accuracy measures of intraindividual variability predict cognitive performance in Alzheimer's disease. *Front. Hum. Neurosci.* **12**, 124.
26. Gogtay N *et al.* 2004 Dynamic mapping of human cortical development during

childhood through early adulthood. *Proc. Natl. Acad. Sci.* **101**, 8174–8179.

27. Burciu RG, Ofori E, Archer DB, Wu SS, Pasternak O, McFarland NR, Okun MS, Vaillancourt DE. 2017 Progression marker of Parkinson's disease: A 4-year multi-site imaging study. *Brain* **140**, 2183–2192.
28. Ingalhalikar M *et al.* 2014 Sex differences in the structural connectome of the human brain. *Proc. Natl. Acad. Sci.* **111**, 823–828.
29. Tononi G, Edelman GM. 1998 Consciousness and Complexity. *Science* **282**, 1846–1851.
30. Casey BJ, Giedd JN, Thomas KM. 2000 Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* **54**, 241–257.
31. Garrett DD, Lindenberger U, Hoge RD, Gauthier CJ. 2017 Age differences in brain signal variability are robust to multiple vascular controls. *Sci. Rep.* **7**, 1–13.
32. Yankner BA, Lu T, Loerch P. 2008 The aging brain. *Annu. Rev. Pathol. Mech. Dis.* **3**, 41–66.
33. Zhang D, Shen D. 2012 Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS One* **7**, e33182.
34. Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS. 1995 Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn. Reson. Med.* **34**, 537–541.
35. Bullmore E, Sporns O. 2009 Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198.
36. DeCasien AR, Sherwood CC, Schapiro SJ, Higham JP. 2020 Greater variability in chimpanzee (Pan troglodytes) brain structure among males. *Proc. R. Soc. B* **287**, p.20192858.
37. Rogachov A, Cheng JC, Erpelding N, Hemington KS, Crawley AP, Davis KD. 2016 Regional brain signal variability: A novel indicator of pain sensitivity and coping. *Pain* **157**, 2483–2492.
38. Power JD *et al.* 2011 Functional Network Organization of the Human Brain. *Neuron* **72**, 665–678.
39. Gabrieli JDE, Ghosh SS, Whitfield-Gabrieli S. 2015 Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*. **85**, 11–26.

40. Reinen JM *et al.* 2018 The human cortex possesses a reconfigurable dynamic network architecture that is disrupted in psychosis. *Nat. Commun.* **9**, 1–15.
41. Chén OY, Cao H, Reinen JM, Qian T, Gou J, Phan H, De Vos M, Cannon TD. 2019 Resting-state brain information flow predicts cognitive flexibility in humans. *Sci. Rep.* **9**, 1–16.
42. Chén OY, Crainiceanu C, Ogburn EL, Caffo BS, Wager TD, Lindquist MA. 2018 High-dimensional multivariate mediation with application to neuroimaging data. *Biostatistics* **19**, 121–136.
43. Koban L, Jepma M, López-Solà M, Wager TD. 2019 Different brain networks mediate the effects of social and conditioned expectations on pain. *Nat. Commun.* **10**, 1–13.
44. Diggle P, Heagerty P, Liang K-Y, Zeger S. 2013 *Analysis of Longitudinal Data*. 2nd edn. Oxford, UK: Oxford University Press.
45. Tinbergen N. 1942 *An objectivist study of the innate behaviour of animals*. Leiden, Netherlands: E.J. Brill.
46. Tinbergen N. 1951 *The Study of Instinct*. Oxford, UK: Oxford University Press.
47. Lorenz KZ. 1957 The Nature of Instinct. In *Instinctive Behavior: The development of a modern concept* (ed CH Schiller), pp. 129–175. New York, USA: International Universities Press.
48. Lorenz KZ, Tinbergen N. 1957 Taxis and Instinct: Taxis and instinctive action in the eggretrieving behavior of the Graylag Goose. In *Instinctive Behavior: The development of a modern concept* (ed HC Schiller), New York, USA: International Universities Press.
49. Griffiths PE. 2004 Instinct in the ’50s: The British reception of Konrad Lorenz’s theory of instinctive behavior. *Biol. Philos.* **19**, 609-631.
50. Brigandt I. 2005 The instinct concept of the early Konrad Lorenz. *J. Hist. Biol.* **38**, 571–608.
51. Burkhardt RW. 2005 *Patterns of Behavior: Konrad Lorenz, Niko Tinbergen and the Founding of Ethology*. Chicago, USA: University of Chicago Press.
52. Griffiths P. 2020 The Distinction Between Innate and Acquired Characteristics. Stanford Encycl. Philos. In *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition; ed Edward N. Zalta).
53. Kricker A, Armstrong BK, English DR. 1994 Sun exposure and non-melanocytic skin cancer. *Cancer Causes Control* **5**, 367–392.
54. Bertoli S, Leone A, Battezzati A. 2015 Human bisphenol a exposure and the

"diabesity phenotype". *Dose-Response 1–12. doi: 10.1177/1559325815599173.*

55. Hook EW, Marra CM. 1992 Acquired syphilis in adults. *N. Engl. J. Med.* **326**,1060–1069.
56. Kant I. 1787 *Kritik der reinen Vernunft, 2nd edition, translated as Critique of pure reason by WS Pluhar*. Indianapolis, USA: Hackett.
57. Pears DF. 1953 *Incompatibilities of colours*. Oxford, UK: Blackwell.
58. Bar-Haim Y, Ziv T, Lamy D, Hodes RM. 2006 Nature and nurture in own-race face processing. *Psychol. Sci.* **17**, 159–163.
59. Waddington CH. 2014 *The strategy of the genes: A discussion of some aspects of theoretical biology*. Abingdon, UK: Routledge.
60. Piper R. 2007 *Extraordinary Animals: An Encyclopedia of Curious and Unusual Animals*. London, UK: Greenwood Press.
61. Quinn AE, Georges A, Sarre SD, Guarino F, Ezaz T, Marshall Graves JA. 1983 Temperature-Dependent Sex Determination in Vertebrates. *Science* **316**, 411.
62. Ge C, Ye J, Weber C, Sun W, Zhang H, Zhou Y, Cai C, Qian G, Capel B. 2018 The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. *Science* **360**, 645–648.
63. Hunte W, Myers RA, Doyle RW. 1985 Bayesian mating decisions in an amphipod, Gammarus lawrencianus Bousfield. *Anim. Behav.* **33**, 366–372.
64. Luttbeg B. 1999 Reproductive decision-making by female peacock wrasses: flexible versus fixed behavioral rules in variable environments. *Behav. Ecol.* **10**, 666– 674.
65. Valone TJ. 1992 Information for patch assessment: A field investigation with black-chinned hummingbirds. *Behav. Ecol.* **3**, 211–222.
66. Valone TJ, Giraldeau LA. 1993 Patch estimation by group foragers: What information is used? *Anim. Behav.* **45**, 721–728.
67. Valone TJ, Brown JS. 1989 Measuring patch assessment abilities of desert granivores. *Ecology* **70**, 1800–1810.
68. Sohn H, Narain D, Meirhaeghe N, Jazayeri M. 2019 Bayesian Computation through Cortical Latent Dynamics. *Neuron* **103**, 934–947.
69. Knill DC, Richards W. 1996 *Perception as Bayesian inference*. Cambridge, UK: Cambridge University Press.
70. Griffiths LT, Kemp C, Tenenbaum B. 2008 Bayesian models of cognition. In

Cambridge Handbook of Computational Cognitive Modeling (ed R Sun), pp. 59–100. Cambridge, UK: Cambridge University Press.

71. Körding KP, Wolpert DM. 2004 Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247.
72. Hubel DH, Wiesel TN. 1968 Receptive fields and functional architecture of monkey striate cortex. *J.Physiol.* **195**, 215–243.
73. Desimone R, Albright TD, Gross CG, Bruce C. 1984 Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* **4**, 2051–2062.
74. Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K. 2013 The WU-Minn Human Connectome Project: An overview. *Neuroimage* **80**, 62–79.
75. Barabási AL, Oltvai ZN. 2004 Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **5**, 101–113.
76. Leclerc RD. 2008 Survival of the sparsest: Robust gene networks are parsimonious. *Mol. Syst. Biol.* **4**, 213.
77. van den Heuvel MP, Sporns O. 2013 Network hubs in the human brain. *Trends Cogn. Sci.* **17**, 683–696.
78. Chén OY. 2019 The Roles of Statistics in Human Neuroscience. *Brain Sci.* **9**, 194.
79. Fan J, Lv J. 2008 Sure independence screening for ultrahigh dimensional feature space. *J. R. Stat. Soc. Ser. B* **70**, 849–911.
80. Ravikumar P, Lafferty J, Liu H, Wasserman L. 2009 Sparse additive models. *J. R. Stat. Soc. Ser. B* **71**, 1009–1030.
81. Chipman HA, George EI, McCulloch RE. 2012 BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **4**, 266–298.
82. Johnstone IM, Titterington DM. 2009 Statistical challenges of high-dimensional data. *Philos. Trans. R. Soc. A* **367**, 4237–4253.
83. Fan J, Liao Y. 2014 Endogeneity in high dimensions. *Ann. Stat.* **42**, 872–917.
84. Liang K-Y, Zeger SL. 1986 Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* **73**, 13–22.
85. Chén OY *et al.* 2020 Thou Shalt Not Reject the *P*-value. *arXiv* **2002.07270**.
86. Miller KL *et al.* 2016 Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536.
87. Huber PJ. 1985 Projection Pursuit. *Ann. Stat.* **13**, 435–475.

88. Aertsen A, Preissl H. 1991 Dynamics of Activity and Connectivity in Physiological Neuronal Networks. *Nonlinear Dyn. Neuronal Networks* **9**, 1303–1350.
89. Friston KJ. 2011 Functional and Effective Connectivity: A Review. *Brain Connect.* **1**, 13–36.
90. Shmueli G. 2010 To explain or to predict? *Stat. Sci.* **25**, 289–310.
91. Friedman J, Hastie T, Tibshirani R. 2001 *The elements of statistical learning*. 2nd edn. New York, USA: Springer series in statistics.
92. Stoker T, Torsney K, Barker R. 2018 Pathological Mechanisms and Clinical Aspects of GBA1 Mutation-Associated Parkinson's Disease. In *Parkinson's Disease: Pathogenesis and Clinical Aspects* (eds S TB, Greenland JC), Brisbane, Australia: Codon Publications.
93. Chén OY, Lipsmeier F, Phan H, Prince J, Taylor K, Gossens C, Lindemann M, de Vos M. 2020 Building a machine-learning framework to remotely assess Parkinson's disease using smartphones. *IEEE Trans Biomed Eng* **67**, 3491–3500.
94. Neyman J. 1935 Statistical problems in agricultural experimentation. *J. R. Stat. Soc.* **2**, 107–180.
95. Rubin DB. 1974 Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701.
96. Rosenbaum PR, Rubin DB. 1983 The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* **70**, 41–55.
97. Dehejia RH, Wahba S. 1999 Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *J. Am. Stat. Assoc.* **94**,1053–1062.
98. Caliendo M, Kopeinig S. 2008 Some practical guidance for the implementation of propensity score matching. *J. Econ. Surv.* **22**, 31–72.
99. Dehejia RH, Wahba S. 2002 Propensity score-matching methods for nonexperimental causal studies. *Rev. Econ. Stat.* **84**, 151–161.
100. Vaux DL, Fidler F, Cumming G. 2012 Replicates and repeats-what is the difference and is it significant? A brief discussion of statistics and experimental design. *EMBO Rep.* **13**, 291–296.
101. Woo CW, Chang LJ, Lindquist MA, Wager TD. 2017 Building better biomarkers: Brain models in translational neuroimaging. *Nat. Neurosci.* **20**, 365–377.

102. Pearl J. 2011 *Causality: Models, reasoning, and inference, second edition*. Cambridge, UK: Cambridge University Press.
103. Pearl J. 1993 *Graphical Models, Causality and Intervention*. http://www.jstor.org/stable/2245965.
104. Pearl J, Robins JM, Greenland S. 1999 Confounding and Collapsibility in Causal Inference. *Stat. Sci.* **14**, 29–46.
105. Hinton G. 2005 What kind of a graphical model is the brain? *Proc. Intl. Jt. Conf. Artif. Intell.* **5**, 1765–1775.
106. Romei V, Thut G, Mok RM, Schyns PG, Driver J. 2012 Causal implication by rhythmic transcranial magnetic stimulation of alpha frequency in feature-based local vs. global attention. *Eur. J. Neurosci.* **35**, 968–974.
107. Lipton RB, Pearlman SH. 2010 Transcranial Magnetic Simulation in the Treatment of Migraine. *Neurotherapeutics* **7**, 204–212.
108. Epstein RJ. 1989 The fall of ols in structural estimation. *Oxf. Econ. Pap.* **41**, 94–107.
109. Stock JH, Trebbi F. 2003 Retrospectives: Who Invented Instrumental Variable Regression? *J. Econ. Perspect.* **17**, 177–194.
110. Bickel PJ, Levina E. 2008 Covariance regularization by thresholding. *Ann. Stat.* **36**, 2577–2604.
111. El Karoui N. 2008 Operator norm consistent estimation of large-dimensional sparse covariance matrices. *Ann. Stat.* **36**, 2717–2756.
112. Tenenbaum JB, De Silva V, Langford JC. 2000 A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323.
113. Roweis ST, Saul LK. 2000 Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**, 2323–2326.
114. Belkin M, Niyogi P. 2003 Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**, 1373–1396.
115. Donoho DL, Grimes C. 2003 Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proc. Natl. Acad. Sci.* **100**, 5591–5596.
116. Johnson WB, Lindenstrauss J. 1984 Extensions of Lipschitz mappings into a Hilbert space. *Contemp. Math.* **26**, 189–206.
117. Pearson K. 1901 LIII. *On lines and planes of closest fit to systems of points in space*. *Philos. Mag. Ser.* **2**, 559–572.
118. Van Der Maaten L, Hinton G. 2008 Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605.