



HAL
open science

Analyzing Brain Networks in Population Neuroscience: A Case for the Bayesian Philosophy

Danilo Bzdok, Dorothea Floris, Andre Marquand

► **To cite this version:**

Danilo Bzdok, Dorothea Floris, Andre Marquand. Analyzing Brain Networks in Population Neuroscience: A Case for the Bayesian Philosophy. *Philosophical Transactions of the Royal Society of London. B (1887–1895)*, 2020. hal-02447507

HAL Id: hal-02447507

<https://hal.science/hal-02447507>

Submitted on 21 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Invited Manuscript for Special Issue:
“Unifying the essential concepts of biological networks”*

Analyzing Brain Networks in Population Neuroscience: A Case for the Bayesian Philosophy

Short title: Bayesian Tools for Big-Data Neuroscience

Danilo Bzdok^{1,2,3}, Dorothea L. Floris⁴, Andre F. Marquand^{4,5}

¹ Department of Biomedical Engineering, McConnell Brain Imaging Centre, Montreal Neurological Institute, Faculty of Medicine, McGill University, Montreal, Canada

² Mila – Quebec Artificial Intelligence Institute, Montreal, Canada

³ Parietal Team, Institut National de Recherche en Informatique et en Automatique (INRIA); Neurospin, Commissariat à l’Energie Atomique (CEA) Saclay, Gif-sur-Yvette, France

⁴ Department of Cognitive Neuroscience, Radboud University Medical Centre, Donders institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands

⁵ Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, U.K.

Keywords: connectome-based prediction, uncertainty, confounding influences, statistical learning

Abstract

Network connectivity fingerprints are among today’s best choices to obtain a faithful sampling of an individual’s brain and cognition. Widely available MRI scanners can provide rich information tapping into network recruitment and reconfiguration that now scales to hundreds and thousands of humans. Here we contemplate advantages of analyzing such connectome profiles using Bayesian strategies. These analysis techniques afford full probability estimates of the studied network coupling phenomena; provide analytical machinery to separate epistemological uncertainty and biological variability in a coherent manner; usher towards avenues to go beyond binary statements on existence vs. non-existence of an effect; and afford credibility estimates around all model parameters at play, which thus enable single-subject subjects predictions with rigorous uncertainty intervals. We illustrate the brittle boundary between healthy and diseased brain circuits by autism spectrum disorder as a recurring theme where, we argue, network-based approaches in neuroscience will require careful probabilistic answers.

Introduction

In network-centered research, as well as many other fields of neuroscience, drawing statistical conclusions from brain data is essential to understand the measurements of the studied phenomenon despite the presence of noise. Typical examples include inferring whether a given functional brain connection is strengthened or weakened by administering a certain environmental stimulus or predicting a clinical diagnosis of a given individual on the basis of neuroimaging data. In this article, we argue that adopting a Bayesian perspective on network-based explanation and modelling offers several benefits, which arise from the ability to coherently handle uncertainty in developing model predictions about phenomena observed in network circuits.

Bayesian analysis and conceptualization has a long history, with origins in the 18th century (Bayes, 1763). In essence, the Bayesian framework treats all parameters in a given model as random variables and quantifies their uncertainty using Bayes rule. For a model \mathcal{M} with parameters θ this principle indicates that the *prior* belief in the probability of the parameters $p(\theta|\mathcal{M})$ should be updated in light of observed data, y , to derive the *posterior* distribution over the entire set of model parameters, $p(\theta|y, \mathcal{M})$.

$$p(\theta|y, \mathcal{M}) = \frac{p(y|\theta, \mathcal{M})p(\theta|\mathcal{M})}{p(y|\mathcal{M})}$$

The term $p(y|\theta, \mathcal{M})$ denotes the *likelihood* and specifies a generative model that describes how the data may have come about. The denominator is referred to as the *marginal likelihood* and is obtained by integrating out the parameters $p(y|\mathcal{M}) = \int p(y|\theta, \mathcal{M})p(\theta|\mathcal{M}) d\theta$. Posterior computation is typically intractable for all but the

simplest models. This is why posterior parameter estimation often invites the use of numerical approximations or sampling methods (Gelman et al., 2014; MacKay, 2003). Ready computability is probably the major hurdle for more widespread adoption of the Bayesian framework in network-based approaches.

It is important to recognise that the Bayesian philosophy of data analysis operates with a deeper and more universal concept of probability than is assumed by most of the quantitative frameworks commonly used in many areas of brain network analysis. In particular, under the frequentist paradigm (Cox, 2006), probability reflects long run frequencies of repeatable events (e.g. ‘the probability of rolling a 6 on this dice is $1/6$ ’). Under the Bayesian paradigm, probabilities reflect degrees of belief in a given proposition (e.g. ‘there is a low probability that the amygdala will increase functional connectivity 100 times more (or less) in autism vs. health’), which may not be repeatable. In a network modelling context, investigators routinely resort to frequentist notions, especially for hypothesis testing against a null distribution, for example to define the probability that a given brain region shows more neural coupling strengths than would be expected under the null hypothesis of baseline activity.

According to the frequentist philosophy, the data-generating mechanism underlying observed network dynamics is fixed and only the observed measurements from biological networks have a probabilistic component. Inference about the model is therefore indirect, quantifying the agreement between the observed biological data and the data generated by a putative model (for example, the null hypothesis). In the Bayesian philosophy, instead, inference quantifies the uncertainty about the data-generating mechanism by the prior distribution and updates it with the data observed from biological networks to obtain the posterior distribution (Figure 1). Inference

about the model is therefore obtained directly as a probability statement provided by the derived posterior parameter distributions.

Over the past decade, the alternative to sampling-based approaches, namely, variational Bayes approaches, dominates in neuroimaging analyses of (effective) connectivity. One example of this approach dynamic causal modelling (DCM) and is a prominent Bayesian method for characterizing imaging time series (Friston et al., 2003; Friston et al., 2014). Crucially, variational approaches avoid the computational cost of sampling by assuming a particular form for the posterior density (Beal, 2003; Fox and Roberts, 2012; Friston et al., 2007; Zhang et al., 2018). This leads to analytic update equations that allow people to perform efficient and quick Bayesian inference on the parameters of their models. Furthermore, variational Bayes considerably finesses the problems of Bayesian model comparison, selection and reduction, to which we will return below.

Direct quantification of uncertainty is the central motif of the Bayesian framework (Gelman et al., 2013). Bayesian modelling aims to coherently incorporate uncertainty throughout the analysis such that uncertainty in the parameter estimation is carefully propagated through the generative model to form predictions about the biological system under study. Mathematical proofs show that probability theory is a unique way that this can be achieved; on the basis of a simple and common-sense set of axioms (Cox, 1946; Jaynes, 2003). In short, any system of reasoning that coherently manages uncertainty for complex biological systems must be consistent with the rules of probability. As such, when carrying out Bayesian analysis of biological networks, the analyst naturally goes beyond point estimates of parameters, such as a value indicating the network connectivity strength between the amygdala and the prefrontal cortex. Rather than a single number (e.g. one Pearson correlation value of $\rho=+0.27$), full probability distributions are placed over all quantities in network modelling which are updated in the light of the brain data at hand. Based on the build Bayesian model, new predictions can be formed for incoming data points by averaging (i.e. integrating over) the joint posterior distribution over all model variables on the table.

A defining characteristic of the Bayesian philosophy is that this modelling regime requires the specification of a prior distribution, reflecting the beliefs about the model parameters before observing any data on brain networks. Thus, each model parameter enters with a fully specified probability distribution, whether or not biological observations have already been brought into play or not. If prior information is available (e.g. the topology of structural connectivity constraining functional connections, which functional connections describe intra-network versus between-network connections or whether connections between subcortical areas may be harder to measure than those between cortical areas), this can be incorporated in the model. This helps guide the parameter updates to biologically plausible ranges in the face of new observations from brain networks, while still permitting solutions that exceed the pre-set ranges to the extent supported by the data. Even in the absence of definite strong, biologically grounded a-priori information, generic priors can be employed to exert a regularizing or smoothing effect on the parameter estimation (e.g. to prevent overfitting connectomic profiles of a subject sample that may not extrapolate well to the broader population). However, the specification of the priors over model parameters is often a point of criticism for Bayesian methods. This is because it can often be difficult to specify informative priors if the number of variables are large or the dependencies between them are complex such as in many biological systems (Woolrich et al., 2009). Moreover, it is often not straightforward to specify priors that convey a lack of prior knowledge (Cox, 2006; Jeffreys, 1946). Nevertheless, it is important to recognise that any network modelling framework is predicated on certain assumptions. The fact that the Bayesian approach forces these to be made explicit by the investigator can be viewed as a strength.

Notions of probability: Methodological uncertainty and biological variability

The type of probability that is actually being modelled is an important epistemological distinction that is often under-appreciated across biological sciences and in network-based research on the brain in particular. Under the Bayesian conceptualization, probabilities can service multiple different purposes in network modelling (Cox, 2006): probability may be treated in a ‘phenomenological’ manner to quantify natural biological variation in the brain data (e.g. how different are amygdala-prefrontal connections across subjects in the population). However, probability can also be framed in an ‘epistemological’ manner to quantify modelling uncertainty in estimating parameter values (e.g. how unsure are we about different amygdala-prefrontal connections due to finite sampling from biological networks).

This consideration reflects the distinction in statistical machine learning (Gal, 2016; Kendall and Gal, 2017) between ‘aleatoric’ uncertainty which reflects inherent variation in the measured phenomenon in biology that cannot be reduced with acquiring more observations of the biological system and ‘epistemic’ uncertainty which reflects uncertainty in our knowledge of model parameters and data densities can be reduced by adding more observations. Unfortunately, this nomenclature confounds the notions of variability and uncertainty described above. To simplify the discussion, we henceforth distinguish between (biological) *variability* and (methodological) *uncertainty*. For completeness, we note that in some cases it may

be desirable to further decompose epistemic uncertainty (e.g. due to scanner noise or interpolation error).

Importantly, most dominant frequentist approaches currently used for brain network modelling conflate variability and uncertainty to a certain degree. Frequentist approaches – at best – provide post-hoc estimates of model *uncertainty* using supplementary techniques such as bootstrapping (Efron, 1979). For most network-focused applications, accurately quantifying *variability* is of primary explanatory interest, while minimizing or properly accounting for *uncertainty*. Indeed, in the physical and life sciences, uncertainty quantification is now regarded as one of the most important estimation challenges (e.g. Lima et al., 2016; Mann et al., 2017; Matta and Massa, 2017). This is especially the case in weather prediction and climate change (Nikam and Meshram, 2013). It is also an important aspect of Bayesian model comparison and the way research hypotheses are tested within a Bayesian framework.

To provide a concrete example from imaging neuroscience, normative modelling is a recently introduced technique that aims to map centiles of variation, such as the functional connectivity strength between amygdala and prefrontal cortex, across a reference cohort in an analogous manner to the use of growth charts in pediatric medicine (Marquand et al., 2019; Marquand et al., 2016a). For example, by plotting biological parameters as a function of age (or other clinically relevant variables), normative modelling enables statistical conclusions as to where the network coupling profile of each individual participant falls within the population range. This modelling tactic can therefore be used to chart *variability* in biological networks relevant to

many disorders including autism and detect the biological signatures of brain disorders in an anomaly detection setting (Zabihi et al., 2018).

In such applications, the primary interest is in modelling inter-individual variation across the cohort whilst accounting for modelling *uncertainty* such as noise intrinsic to the fMRI signals from brain networks. For such neuroscience applications, the ability to jointly model different sources of variation and appreciate uncertainty in the same modelling instance is an important advantage of the Bayesian culture. For example, using Bayesian methods the investigator can use separate variance components to model variation in age-related connection strength across a population cohort and the uncertainty in that estimation, due for example to data sampling density (e.g. fewer female subjects, or less high-functioning patients). In contrast, classical methods may also be used for normative modelling. Confidence intervals for the centiles of variation could be derived using bootstrapping (Huizinga et al., 2018). However, bootstrapped frequentist models cannot easily be used to draw probabilistic conclusions on new, unseen connectomic data. In Bayesian analysis instead the directly estimated posterior distributions qualifying each model parameter can be readily used to form fully probabilistic predictions, such as on an individual's diagnosis for autism on the basis of connectome fingerprints.

The value of Bayesian analysis for the goal of delineating quantities of variability and uncertainty in connectivity analysis has been advertised through a body of literature (e.g. Bowman et al., 2008; Friston et al., 2008; Friston et al., 2002a; Friston and Penny, 2003; Friston et al., 2002b; Penny et al., 2005; Woolrich, 2012; Woolrich et al., 2004; Woolrich et al., 2009). More specifically, in the context of brain networks,

Bayesian methods have been applied for improving the estimation of whole-brain connectivity profiles (Colclough et al., 2018; Hinne et al., 2014) in finding parcellations of different brain networks (Janssen et al., 2015), for causal inference in fMRI (Mumford and Ramsey, 2014) and for multi-modal data fusion (Groves et al., 2011). These existing neuroimaging applications have largely focused on datasets of modest size, for which Bayesian methods are well-suited due to the regularizing effect exerted by the imposed priors and the guidance of parameter updates by existing neuroscience knowledge. As such, generic priors can be used to deprioritize exceedingly large model parameters to discourage unrealistic model parameter estimates. Such smooth bounding of suboptimal parameter candidates during model estimation helps guard against overfitting to seemingly coherent patterns in the connectivity fingerprints of the subjects. In addition to previous applications, we argue here that Bayesian methods also provide an excellent tool for large, population-based cohorts, which are gaining center stage in clinical neuroimaging (Di Martino et al., 2014; Miller et al., 2016b; Smith et al., 2015a; Van Essen et al., 2013; Volkow et al., 2017).

There are several reasons for the suitability of the Bayesian framework in the 'big data' era (cf. Bzdok et al., 2019; Bzdok and Yeo, 2017): the ability to separately quantify variability and generate explainable insight in the natural phenomenon under study and uncertainty in the model under use is likely to be instrumental to understanding inter-individual variations across large cohorts (cf. above). Its importance is increasingly recognised in IQ prediction based on connectivity fingerprints and other successful examples (Finn et al., 2015; Foulkes and Blakemore, 2018; Marquand et al., 2019; Seghier and Price, 2018). Bayesian

methods are also appealing because they provide estimates of the plausible range of a parameter value given the brain data. In contrast, in large samples classical null-hypothesis testing methods can easily reject the null hypothesis for nearly all values (e.g. all network nodes in a classical frequentist connection-wise analysis), even though the underlying effects are of negligible magnitude (see Friston and Penny, 2003). It should be obvious that quantifying methodological *uncertainty* is critical for optimal decision-making in medicine (Bishop, 2006). For example, for predicting an autism diagnosis on the basis of MRI scans, where uncertainty arises at multiple levels: not only in the diagnosis itself (i.e. at the level of clinical presentation), but also at the level of the underlying biology (e.g. the connectivity strength in a network modelling context).

In this paper, we will provide a conceptual overview of the aim and utility of the Bayesian modelling framework in clinical neuroscience, focusing on the use of such methods for generating explainable insights on connectomics. Functional connectivity fingerprints are particularly valuable for capturing salient characteristics of momentary states of conscious awareness and for predicting individual differences in cognition (Finn et al., 2015; Rosenberg et al., 2016; Smith et al., 2015a). These analytical techniques are widely applicable to predicting symptomatology across many clinical populations (Fornito et al., 2015; Xia et al., 2018).

Hierarchical Bayesian modelling: Appreciating covariates of population stratification

The boundary between signal and noise is often hard to identify; let alone to know prior to data analysis. It is common practice in many empirical sciences, including network analysis in imaging neuroscience (e.g. Miller et al., 2016a; Smith et al., 2015b), to adjust for nuisance variance in the data in two separate steps. In a first modelling step, variation that can be explained by nuisance covariates is removed, typically using linear-regression-based deconfounding. In a subsequent modelling step, the remaining variation in the data is then fed into the actual statistical model of interest used to draw neuroscientific conclusions on brain network phenomena. As such, the final explanation is typically grounded in model parameter estimates from a version of the original data, in which any linear association with the considered nuisance covariates, such as age- and sex-related differences between individuals, has already been comprehensively removed beforehand. In this approach to network modelling, the implicit but critical assumption is that any target effects of interest in the brain data, such as for the goal of classifying neurotypicals from individuals with a diagnosis of autism based on connectomic fingerprints, is treated largely separately of what is measured by the nuisance covariates.

In many brain disorders, including autism, the distinction between signal and noise may be more ambiguous than established analysis workflows belie. Age, sex, and motion are routinely chosen as nuisance covariates. However, the majority of autism samples include 3-5 times more males who carry a diagnosis of autism than females (Kanner, 1943; Lai et al., 2017b; Scott et al., 2002), reflecting differences in

prevalence in the wider population. While several reasons can be brought forward (Floris et al., 2018; Goldman, 2013; Lai et al., 2015; Schaafsma and Pfaff, 2014), it has been speculated that the discrepant prevalence of autism may point to a more profound distinction in the etiology of the disease, potentially linked to its triggering life events, underlying pathophysiological mechanisms, and ensuing coping strategies. Preceding removal of sex-related signal in the data can also remove information of and preclude insightful explanations about sex-specific disease pathways in autism or lead to spurious findings or incorrect conclusion (Miller and Chapman, 2001; Miller et al., 2016a). Let's consider a hypothetical scenario where amygdala-prefrontal connectivity is pathologically increased in male patients, but pathologically decreased in female patients. Here, a preceding deconfounding step for sex would largely remove this sex-dependent aspect, which, however, truly is a characteristic of disease biology, from subsequent statistical analysis and scientific conclusion.

In a similar spirit, the age trajectories of male and female individuals with autism, including the manifestations in underlying network biology, may be different in multiple ways. For instance, a commonly described clinical feature of autism is that females are more often diagnosed later in life (Schaafsma and Pfaff, 2014). Better coping strategies and more successful camouflaging behavior in women with autism is a common explanation for this age-related divergence (Lai et al., 2017a). Consequently, removing age-related variance in brain network measurements as a "data preprocessing" routine, expected by peer scientists and paper reviewers in the community, may systematically withhold insights that can teach us something about the age-dependent development of autism in different strata of the population. We

therefore argue that it is often more pertinent to model shared variance explicitly, such as in jointly modelling age-dependent connectivity variation and autism-dependent variation in the functional connectome, for which the Bayesian framework is well suited. For example, Bayesian analysis could answer a question such as “How certain are we that amygdala-prefrontal connectivity strength is similar or dissimilar in certain subgroups, such as when stratifying by sex or lifespan?”.

Rather than resorting to a deterministic decision in a black-or-white fashion, Bayesian hierarchical modelling (BHM) is a natural opportunity to quantify the separate contributions by answering which sex-, age- and motion-related components in network connectivity couplings are related to autism-related model parameters with which magnitude and how certain the investigator can be about it. A set of sources of variation in the brain data can be directly integrated in a single model estimation, instead of carrying out initial confound and later effect analyses (cf. above). Said in yet another way, BHM allows for explicit modelling of the group differences in functional brain connections in disease vs. control groups as linked to the question of how much any group difference are influenced by age, sex, and motion variation in the functional connectivity data by hierarchically accounting for dependencies between them.

Removing age, sex and motion related information from the data in an isolated modelling step hides important information that can be instrumental in guiding the parameter estimation of the model actually used to gain biological insight. While this goal can also be accommodated in a non-Bayesian setting (e.g. using linear mixed effects models), the Bayesian formulation is appealing because it coherently

propagates uncertainty through different levels of the model and can therefore more readily disentangle different sources of variability and uncertainty.

One prevalent form of Bayesian hierarchical modelling is known as parametric empirical Bayes (PEB) (Friston et al., 2016; Kass and Steffey, 1989). This now underwrites most of the between subject analyses using dynamic causal modelling. In brief, parametric empirical Bayes rests on a hierarchical Bayesian model in which random effects at the within and between subject level can be accommodated. Usually, these hierarchical models are based on a general linear models at higher levels; hence parametric. Below, we will consider nonparametric empirical Bayesian models.

Motion during brain scanning is one of the measurements that is widely used to remove variation from connectivity data; before starting the actual functional connectivity analysis. A few years ago, neuroimaging investigators have reported a seemingly distinctive pattern of maturing functional network fluctuations with weakening short-range and growing long-range connections that slowly change during child development (Dinstein et al., 2011; Dosenbach et al., 2010). Investigators speculated these findings to mean that normal children start life with prominent short-range connectivity, which then weakens over the life span in healthy controls; vice versa for long-range connectivity (Belmonte et al., 2004). Individuals carrying a diagnosis of autism were then found to show more short-range and less long-range connectivity links (Geschwind and Levitt, 2007; Keary et al., 2009), especially in children. Unfortunately, it later became apparent that excessive head movements during brain scan acquisition reliably entailed artefacts with these same

connectivity patterns in functional brain connectivity, previously thought to reflect impaired brain maturation (Power et al., 2012; Van Dijk et al., 2012), which entailed several retractions of high-profile papers (<https://www.spectrumnews.org/news/movement-during-brain-scans-may-lead-to-spurious-patterns/>).

On the other hand, at the behavioral level, it is well established that people with a diagnosis of autism exhibit greater degrees of movement than healthy controls (Nordahl et al., 2008; Yendiki et al., 2014). As such, unusually high body movement can be argued to be a hallmark feature of autism, but is now recognized to also be a reason for spurious functional connectivity findings. Put differently, it is hard to give a single clear-cut answer which aspects of functional connectivity signals corresponds to motion-related noise and which aspects corresponds to biologically informative signals in functional connectivity synchronization between brain regions. With and without a given adjustment for motion-related influences, distinguishing functional connectivity fingerprints in autism reflect different statistical questions (Pearl and Mackenzie, 2018). These data modelling scenarios correspond to two equally valid questions depending on the scientific purpose. Adjustment relates to partitioning a population into groups that are homogenous according to the de-confounding variable - there may be no single right or wrong. Bayesian analysis can help in quantifying uncertainty via joint probability distributions that explicitly incorporate how body motion measurements are related to network connectivity strengths and to other measurements of interest in an integrated approach.

The inferential grip of insights about brain network coupling can thus be enhanced by findings with models acknowledging variation at different scales. In this way, BHM allows asking more ambitious questions using hierarchical population models of brain connectomics in strata of individuals. Young people with autism are different from old people with autism as reflected in their connectome profiles. An additional and not mutually exclusive source of variation is that male autism is different from female autism, conjointly across lifespan. We can estimate differences in network connectivity between autism and control groups by modelling hierarchical dependencies between multiple sources including covariates, like age, sex, and motion, with parameters corresponding to network connectivity measurements. This multi-level modelling setup allows for partial pooling of information between measurements suspect to exert confounding influence and genuine measurements of brain signals. For instance, neuroscientists may find that increased amygdala-prefrontal connectivity in autism is particularly characteristic for females who are in early childhood and tend to move their head little in the brain scanner as part of a joint posterior parameter distribution incorporating all measured sources of variability. Additionally, sex imbalance is often encountered in population samples of autism which can reflect the population prevalence or explicit exclusion of female cases. Imbalance in the considered participants in each group can be explicitly handled by BHM, with appropriate accounting for uncertainty. To adjust for these differences in naturally occurring group size we can avoid being misled in the way common single-level models typically would be. As such the often made a-priori distinction into signal and noise, as a separate preprocessing step, can be relaxed by combining and integrating statistical evidence from disparate sources in a single probabilistic model estimation (Efron and Hastie, 2016).

As a more concrete scenario for a key strength of flexible Bayesian hierarchical approaches, age, sex and motion may be modeled in well-defined nested relationships to make predictions from brain connectomic variation. It is conceivable that the tendency for body motion in the scanner is a function of age and differs by sex. As such, the investigator may wish to specify a generative model, where male and female participant prior distributions are at the top level, from which probabilistic distributions for age decades are sampled that, in turn, give rise to the (age/sex-dependent) motion covariate distributions. During model estimation, partial pooling between the age-, sex-, and motion dependencies calibrate parameter shrinkage in a data-dependent fashion to achieve optimal prediction performance. The obtained posterior parameter distributions then allow the investigator to draw careful conclusions about the multi-level relationships between effects from several inter-dependent sources of age, sex, and motion variation in how they relate to connectome-based predictions. Importantly, such joint modeling of sources of randomness is challenging in classical (non-Bayesian) general linear models.

The importance of saying no: Uncertainty estimates for single-subject predictions

As one of various supporting hints for the biological basis of autism, the integrity of the amygdala in the limbic network was repeatedly highlighted to differ in patients with autism, which is thought to play a role in impaired social interaction (Baron-Cohen et al., 2000). Statistically significant differences in the amygdala in autism led

to varying reports in different patient samples (Kim et al., 2010; Nacewicz et al., 2006). Thus, this disease manifestation does not appear to be present in every single autism patient, nor is it consistently present on average in every patient sample recruited for studies that compare healthy and diagnosed individuals. Asking whether or not a strict categorical difference exists in a specific brain region in individuals on the autism spectrum may simply be a suboptimal analytical approach for the job.

Put differently, any modelling technique that is designed to give categorical black-and-white answers may be inappropriate for probing disease features that are a) present in autism patients to varying degrees (i.e. reflecting biological *variability*), b) difficult to detect from the noisy behavioral and/or functional connectivity measurements that are available (i.e. reflecting epistemological *uncertainty*), or both. If these two sources of variation have played a role in amygdala studies in autism then using analysis approaches that can only make raw statements declaring presence or absence of an effect may be inherently ill-suited (Kruschke and Liddell, 2018). If the network phenomenon under study is highly variable across people and/or tricky to quantify methodologically, then investigators in one lab may conclude on presence of a difference in connectivity between amygdala and prefrontal cortex on their sample, while another research group studying a different patient sample with the same research question may conclude on absence of group-related connectivity differences. While the answer is seemingly certain in each of these studies, the uncertainty in whether or not an amygdala effect is present in the limbic network of autists comes out at another end (Nosek et al., 2015): lacking reproducibility across different studies that have carried out a dichotomic test for

statistically significant vs. insignificant amygdala alterations in autism samples (He et al., 2019).

In such cases, the conventional frequentist 95% confidence intervals are not the solution that many investigators desire. It is common to hear that a 95% confidence interval means that there is a probability 0.95 that the true parameter value lies within the interval; that is, that we do not have enough evidence to reject the null hypothesis of equal amygdala volume in both groups. In non-Bayesian statistical hypothesis testing, such a statement is never correct, because strict non-Bayesian inference forbids using probability to measure uncertainty about parameters like a measure of amygdala connectivity in healthy vs. diseased individual. Instead, one should say that if we repeated the study and analysis a large number of different samples, then 95% of the computed intervals would contain the true parameter value. The classical 95% confidence interval only takes its meaning in the hypothetical long run of repeatedly analyzing always new samples of controls and patients. Then we expect to be mistaken about presence or absence of amygdala effect in only 5%, that is 1 in 20 of the conducted network connectivity studies.

A particularly clear illustration of this point is Lindley's paradox (https://en.wikipedia.org/wiki/Lindley%27s_paradox). It describes a situation in which a classical statistical analysis suggests a very significant effect, despite the fact that the Bayesian model evidence for the null hypothesis or model is far greater than the alternative. This paradox explains the dangers of over-powered studies that can become too sensitive to trivial effect sizes, while, in a Bayesian setting, would provide evidence for the null hypothesis.

Rather than forcing definitive answers on presence against absence of subtle amygdala effects using null-hypothesis statistical significance testing, Bayesian analysis fully embraces unavoidable variation as an integral part of model building, estimation, and interpretation (Gelman et al., 2014). In the Bayesian paradigm, each component of the model has a fully specified probability distribution; before and after seeing the brain data. As a consequence, a Bayesian model estimating differences in amygdala connectivity in healthy vs. autistic individuals naturally provides estimates of the *degree of difference* at the phenomenological level as well as estimates of the *modelling uncertainty* at the epistemological level. Any amount of divergence between 0 ('no difference') and 1 ('difference') is a possible, legitimate, and interpretable result in the Bayesian posterior parameter estimate, while fully accounting for uncertainty in the parameter estimation.

In this way, the Bayesian modelling regime offers rigorous statements on how much an explanation on a given group difference in the amygdala is justified in the patient sample at hand. The width of the corresponding parameter posterior estimate can be narrow to indicate high certainty in the obtained group difference (cf. Figure 1). In contrast, the posterior distribution can be widely spread out to indicate low methodological certainty and thus limited neuroscientific trustworthiness of the found parameter value reflecting amygdala coupling difference. Taken together, Bayesian modelling directly provides a confidence judgment about each quantity on the table. For example, it allows statements such as: 'under the model, there is a 95% probability that amygdala connectivity to other networks differs between individuals with autism and controls'. If the evidence for the tested difference is ambiguous, we

want this to be the result of the analysis so that we can align the strength of our conclusions with the certainty that the model can afford.

Most modelling approaches following the frequentist philosophy have a harder time telling the investigator when the modelling result is unsure or not. For example, linear support vector classifiers or linear discriminant analysis can be applied to connectomic brain data to vote for autism, rather than control, based on a brittle 51% or a solid 98% probability for evidence of group difference in the amygdala. In other words, Bayesian analysis frameworks are a rare opportunity where the resulting model solution “knows when it does not know”. Moreover, in null-hypothesis statistical testing, the probability of detecting an effect (i.e. statistical power) increases with increasing sample size, even though the effect size (e.g. in terms of group differences in a point estimate for a given parameter) does not (Wagenmakers et al., 2008). Bayesian modelling does not suffer from this shortcoming for the reasons we have outlined above.

Another key benefit of being able to say ‘None’ is that one can compare the evidence for models with and without particular parameters. This affords a very simple form of Bayesian model selection or structure learning; namely Bayesian model reduction (https://en.wikipedia.org/wiki/Bayesian_model_reduction). Using the variational procedures mentioned above, this leads to fast schemes for comparing thousands of models in which various combinations of parameters are “turn off” with appropriate priors.

The ability to say ‘None’ when the investigator asks for whether a group difference either exists or not will probably turn out to be crucial in our efforts towards precision medicine (Arbabshirani et al., 2017; Bzdok and Meyer-Lindenberg, 2018; Stephan et al., 2017). As Bayesian models are fully probabilistic by construction, brain data from a new incoming individual, such as brain scanning yielding amygdala connectivity measures, can be propagated through the already-built model into a *probabilistic prediction* for a single individual at hand. Adding such information can be crucial in a variety of settings in neuroscientific research and clinical practice. First, generating single-subject predictions in a patient may yield different levels of certainty in assessments of autism symptoms related to language, motor behavior, IQ, or social interaction capacities. For example, individuals who are confidently classified may have more severe symptoms in a particular domain, whilst others that are less confidently classified may be more mildly affected. Separate judgments on the certainty of predictive conclusions in each of these symptom domains may turn out to characterize different types of autism in the spectrum, such as high-functioning autism. Second, along the life trajectory, different symptom dimensions of autism may turn out to be predictable based on brain network measurements with higher or lower confidence, which may turn out to be characteristic for developmental periods in autism, or specific for atypical cases or different subtypes of autism. For instance, in women with autism, typically better camouflaging of social deficits (Dean et al., 2017; Hull et al., 2017; Lai et al., 2017a) may lead to social impairment predictions that have non-identical confidence in men with autism. Third, uncertainty is undoubtedly a key asset of treatment response prediction to choose therapeutic interventions tailored to single individuals. In this context, models predicting which treatment option to choose based on an individual’s connectomic profile will be all

the more useful in clinical practice, if such algorithmic recommendations also carry forward information on the forecasting confidence.

Disease subtyping: Towards probabilistic intermediate phenotype discovery

A key challenge in the study of most psychiatric disorders, including autism, is that individuals with the same clinical diagnosis vary considerably from one another in terms of clinical phenotype and underlying network biology. This has led to some investigators proposing that it may be preferable to consider the ‘autisms’ (Geschwind and Levitt, 2007). Many studies have aimed to dissect the clinical phenotype of autism (e.g. Kernbach et al., 2018; Wolfers et al., 2019b), for which functional connectivity provides promising candidate features (e.g. (Easson et al., 2019)). Moreover, since atypicalities are often complex and multifaceted, the features used for this purpose are often high-dimensional (e.g. consider even more whole-brain nodes in functional connectivity matrices) and/or multimodal (e.g. combine measures derived from structural and functional connectivity).

In general, the goal of such connectome-based studies is to derive the latent structure underlying the clinical phenotype (e.g. partitioning the cohort into subtypes) on the basis of psychometric or biological variables, whilst accounting for nuisance variation. There are many ways that this can be achieved, including classical clustering techniques and matrix factorization techniques such as non-negative matrix factorization (NMF) and independent components analysis (ICA). Briefly,

clustering approaches focus on finding subtype clusters in the data, whereas matrix factorization approaches focus on finding useful decompositions of a data matrix under various assumptions. This can be used, for example, to find latent factors that may overlap across individuals in that any given individual may express multiple latent factors (Tang et al., 2019). Whilst these approaches are widely applied in a classical frequentist context, Bayesian variants have also been developed. In addition, highly promising Bayesian ‘non-parametric’ clustering and matrix factorization approaches have been developed such as Dirichlet process mixtures (‘Chinese restaurant processes’) and the ‘Indian buffet’ process (IBP). Adopting a Bayesian approach to such problems confers many benefits, including providing good control over latent representations of the data, thereby helping to attenuate problems with high-dimensional estimation, providing predictive intervals around parameter estimates and predictions and providing flexible noise models for different forms of data. Moreover, Bayesian models are always generative in the sense that they always provide a model for how the brain connectivity measurements may have been generated. Collectively such Bayesian approaches are increasingly applied in clinical and neuroimaging contexts (Groves et al., 2011; Janssen et al., 2015; Ruiz et al., 2014; Schmidt et al., 2009)

A key problem in most classical stratification techniques is the issue of model order selection, or in other words, determining the optimal number of clusters or latent factors for the data at hand (Eickhoff et al., 2015). For example, ‘How many subtypes of autism can be distinguished in a given clinical dataset based on brain connectivity profiles?’. This is a notoriously difficult problem in classical statistics for which no uniquely optimal solution has imposed itself (Bzdok, 2017; Kleinberg, 2002), leading

to suggestions that model order selection is perhaps sometimes largely a matter of taste (Hastie et al., 2009). There are many heuristic approaches for this problem, but these are subject to difficulty in practice. Additionally, choosing between a variety of viable model order selection criteria, which are by themselves objective, still amounts to taking a subjective choice on the number of latent factors best supported by the data. Different *cluster validity criteria* often give different answers or do not indicate a clear preference for one model order over others (Shalev-Shwartz and Ben-David, 2014), nor whether a given clustering solution explains the data better than a continuous model (i.e. with no clusters (Liu et al., 2008)). This has contributed to inconsistencies in the clinical stratification literature such that there are no consistently reported subtypes for autism (Wolfers et al., 2019a) or indeed any psychiatric disorder, despite decades of effort (Marquand et al., 2016b).

Above, we considered parametric empirical Bayesian models (Kass and Steffey, 1989) as prevalent examples of hierarchical Bayesian modelling. The use of Bayesian model comparison and reduction to prune these models provides an efficient way to test hypotheses about the role of any particular brain region affect all connection. A similar functionality can be afforded by Bayesian nonparametric approaches. These provide an appealing solution to this problem because they can automatically adjust the model complexity (e.g. number of clusters or latent factors) on the basis of the data at hand. In other words, non-parametric models allow the flexibility to grow with the number of data points used for model building. The simplest examples of Bayesian non-parametric models are Gaussian process models (Rasmussen and Williams, 2006), which are widely used for non-linear regression and have been used in normative modelling approaches described

above. In a similar manner, Dirichlet process mixtures (DPM) (Ferguson, 1973; Neal, 1992) and ‘Indian buffet’ processes (Griffiths and Ghahramani, 2011) provide an elegant potential solution to the problem of model order selection in clustering and matrix factorization, respectively. For example, the DPM model can be viewed as a clustering model with where the number of clusters is bounded only by the sample size, effectively making the DPM an infinite mixture model (Rasmussen, 2000). This has already been shown to be useful in a recent neuroimaging study on autism (Kernbach et al., 2018). As noted, a very appealing feature of this model is that it is self-calibrating in that it allows the optimal model order (i.e. the number of clusters) to be automatically derived from the data whilst allowing the model order to grow with more data (i.e. increasing representational capacity). In practice, the number of clusters often grows sub-linearly with the number of observations (Neal, 1992). At the same time, by computing (or approximating) the full posterior distribution over the model parameters, this approach helps to attenuate overfitting. This non-parametric clustering approach has clearly desirable features for the stratification of psychiatric disorders such as autism in large data cohorts. Particularly as the size of the available datasets grows (e.g. through larger consortia), such models offer the ability to offer increasingly more fine-grained fractionations of the clinical phenotype. Similarly, in the context of brain networks, this approach has been shown to be useful for automatically parcellating brain networks into component regions (Janssen et al., 2015).

A more recent addition to the Bayesian non-parametric toolbox is the IBP (Griffiths and Ghahramani, 2011) is. The name is derived by analogy to the ‘Chinese restaurant process’ formulation of Dirichlet process mixtures (see (Aldous, 1985)).

The IBP differs in that it does not assume that a single class is responsible for generating each data point (i.e. it does not provide a hard clustering solution). Rather, it allows each data point to express multiple features simultaneously, potentially reflecting multiple causes. Whilst this approach is yet to see extensive applications in brain connectomics, IBPs have been shown to provide an elegant way to model comorbidity in psychiatric disorders, where each individual expresses multiple latent factors to varying degrees (Ruiz et al., 2014).

The key advantage of Bayesian techniques for model order selection in brain network modelling in health and disease is that they provide a formal framework for reasoning over model structures and deducing the plausibility of different candidate structures in view of the data at hand (see e.g. (Ghahramani, 2015; Tenenbaum et al., 2011)).

Conclusion

In this conceptual overview, we provided a pitch for the Bayesian perspective on modelling biological network circuits in the healthy brain and their perturbation in autism spectrum disorder. There are many different ways in which adopting a Bayesian philosophy to analysis and interpretation can open new windows of explanation for neuroscience investigators building on population neuroscience initiatives like UK Biobank, CamCAN, or the Human Connectome Project.

These analysis techniques provide an appealing interpretation of probability in terms of degrees of belief in a proposition, which is more general than the more restricted notion of reasoning about long-run frequencies of repeatable events; provide analytical machinery to separate (methodological) uncertainty and (biological) variability along with a calculus for reasoning about both in a coherent manner; and usher towards avenues away from classical null-hypothesis significance testing, which is particularly valuable in data richness, and may contribute to overcoming the current reproducibility crisis in biomedicine. Finally, Bayesian methods afford estimates of uncertainty around all model parameters at play and can hence form predictions about single individuals by appropriate handling of all considered sources of variation in network approaches. Their value resides for instance in explaining brain network connectivity at different hierarchical scales in the same modelling instance. Promoting such approaches to uncovering key features of biological networks can bear further advantages in the context of data fusion, individual prediction, subgroup stratification of cohorts and for precisely quantifying statistical differences between experimental cohorts. We anticipate that in the coming years,

the Bayesian arsenal will be endorsed for applications in network-focused neuroscience studies more than is currently the norm.

Ethics statement

Not applicable

Competing interests statement

The author declare no conflict of interest

Authors contribution

All three authors wrote and edited the manuscript.

Figure 1

Bayesian model estimation of drug response in 4 autism populations based on whole-brain connectivity profiles. Suppose that a new candidate treatment for autism is being developed, such as a special psychotherapeutic intervention. The investigator now wishes to know whether there is a difference in how reliably a favorable treatment response can be estimated in 4 different subgroups of autism (e.g. these could correspond to male and female high-functioning and not high-functioning individuals carrying a diagnosis of autism) based on inter-subject differences in connectomic fingerprints. Before acquiring any network coupling measurements of patients about to undergo the new treatment option, the investigators pre-suppose that each of the 4 autism subgroups should be expected to have the exact same chance of turning out to be a treatment responder. This initial belief is reflected in 4 probability distributions with equal height (i.e. mean) and equal dispersion (i.e. variance) (*upper panel*) – incorporating fully probabilistic expectations even before any real-world data are considered. After specifying this assumed prior knowledge of equal response probabilities, the Bayesian model is updated (‘parameter updating’) by simultaneously integrating the observed clinical evidence collected from 4 different subgroups of autism patients to achieve a compromise between data-independent prior (*upper panel*) and data-dependent experimental outcomes (*lower panel*). In this example, the consequence is that the prior distributions are carefully adapted in shape – affecting both magnitude and uncertainty – for each autism subgroup. Importantly, the prior probability distribution of showing a favorable response to the novel therapy is re-calibrated in a subgroup-sensitive fashion. After conditioning the model on the actual clinical observations, subgroup group 2 turns out to show the highest posterior parameter distribution. This model solution indicates a strongest chance for the treatment to be successful, relative to the other 3 considered patient subgroups. At the same time, this subgroup posterior parameter distribution features the smallest posterior variance (i.e. highest precision). The narrow dispersion of the posterior treatment effect of subgroup 2 indicates that the investigator can be more sure that the true treatment response probability is close to the estimated treatment effect (i.e. parameter mean). In stark contrast, subgroup 1 shows the widest posterior distribution, which makes explicit that the investigators should be most careful about this estimated therapy response probability. That is, we do have a specific treatment responsiveness for subgroup 1 – in form a concrete number –, the interval of this posterior parameter distribution however also tells us that a much higher or much lower probabilities is quite plausible as well, which is why the obtained parameter mean should be interpreted with caution. Note that subgroups 2 and 3 are predicted to show higher treatment response probability based on connectivity profiles than assumed under the uniform prior of equal response potential with possible implications for clinical practice, rather than succumbing to the dichotomic statement that only the posterior response probability of group 2 is significant and worthy of being reported. Moreover, the conclusion of two subgroups showing evidence for treatment benefits, yet to different degrees with different uncertainty, illustrates the important advantage of Bayesian analysis to allow for fully probabilistic claims in population neuroscience studies. Reproduced with permission (Kruschke, 2011, page 21).

References

- Aldous, D., 1985. Exchangeability and related topics. *École d'Été de Probabilités de Saint-Flour XIII-1983*. Springer, pp. 1-198.
- Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage* 145, 137-165.
- Baron-Cohen, S., Ring, H.A., Bullmore, E.T., Wheelwright, S., Ashwin, C., Williams, S., 2000. The amygdala theory of autism. *Neuroscience & Biobehavioral Reviews* 24, 355-364.
- Bayes, T., Price, N., 1763. An essay towards solving a problem in the doctrine of chances, by the late Rev. Mr. Bayes F.R.S. communicated by Mr. Price, in a letter to John Caton. *Philosophical Transactions of the Royal Society of London* 53, 370-418.
- Beal, M.J., 2003. Variational algorithms for approximate Bayesian inference. university of London London.
- Belmonte, M.K., Allen, G., Beckel-Mitchener, A., Boulanger, L.M., Carper, R.A., Webb, S.J., 2004. Autism and abnormal development of brain connectivity. *Journal of Neuroscience* 24, 9228-9231.
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Springer, Heidelberg.
- Bowman, F.D., Caffo, B., Bassett, S.S., Kilts, C., 2008. A Bayesian hierarchical framework for spatial modeling of fMRI data. *Neuroimage* 39, 146-156.
- Bzdok, D., 2017. Classical statistics and statistical learning in imaging neuroscience. *Front Neurosci* 11, 543.
- Bzdok, D., Meyer-Lindenberg, A., 2018. Machine learning for precision psychiatry: Opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 3, 223-230.
- Bzdok, D., Nichols, T.E., Smith, S.M., 2019. Towards Algorithmic Analytics for Large-Scale Datasets. *Nature Machine Intelligence*, in press.
- Bzdok, D., Yeo, B.T.T., 2017. Inference in the age of big data: Future perspectives on neuroscience. *Neuroimage* 155, 549-564.
- Colclough, G.L., Woolrich, M.W., Harrison, S.J., Lopez, P.A.R., Valdes-Sosa, P.A., Smith, S.M., 2018. Multi-subject hierarchical inverse covariance modelling improves estimation of functional brain networks. *Neuroimage* 178, 370-384.
- Cox, D.R., 2006. *Frequentist and Bayesian Statistics: A Critique*. *Transactions in Particle Physics, Astrophysics and Cosmology*.
- Cox, R.T., 1946. PROBABILITY, FREQUENCY AND REASONABLE EXPECTATION. *American Journal of Physics* 14, 1-13.
- Dean, M., Harwood, R., Kasari, C., 2017. The art of camouflage: Gender differences in the social behaviors of girls and boys with autism spectrum disorder. *Autism* 21, 678-689.
- Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D.A., Gallagher, L., Kennedy, D.P., Keown, C.L., Keyser, C., Lainhart, J.E., Lord, C., Luna, B., Menon, V., Minshew, N.J., Monk, C.S., Mueller, S., Muller, R.A., Nebel, M.B., Nigg, J.T., O'Hearn, K., Pelphrey, K.A., Peltier, S.J., Rudie, J.D., Sunaert, S., Thioux, M., Tyszka, J.M., Uddin, L.Q., Verhoeven, J.S., Wenderoth, N., Wiggins, J.L., Mostofsky, S.H., Milham, M.P., 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry* 19, 659-667.
- Dinstein, I., Pierce, K., Eyster, L., Solso, S., Malach, R., Behrmann, M., Courchesne, E., 2011. Disrupted neural synchronization in toddlers with autism. *Neuron* 70, 1218-1225.

Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett, J.R., Jr., Barch, D.M., Petersen, S.E., Schlaggar, B.L., 2010. Prediction of individual brain maturity using fMRI. *Science* 329, 1358-1361.

Easson, A.K., Fatima, Z., McIntosh, A.R., 2019. Functional connectivity-based subtypes of individuals with and without autism spectrum disorder. *Network Neuroscience* 3, 1-19.

Efron, B., 1979. 1977 RIETZ LECTURE - BOOTSTRAP METHODS - ANOTHER LOOK AT THE JACKKNIFE. *Annals of Statistics* 7, 1-26.

Efron, B., Hastie, T., 2016. *Computer-Age Statistical Inference*. Cambridge University Press.

Eickhoff, S.B., Thirion, B., Varoquaux, G., Bzdok, D., 2015. Connectivity-based parcellation: Critique and implications. *Hum Brain Mapp*.

Ferguson, T.S., 1973. A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics* 1, 209-230.

Finn, E.S., Shen, X.L., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience* 18, 1664-1671.

Floris, D.L., Lai, M.-C., Nath, T., Milham, M.P., Di Martino, A., 2018. Network-specific sex differentiation of intrinsic brain function in males with autism. *Mol Autism* 9, 17.

Fornito, A., Zalesky, A., Breakspear, M., 2015. The connectomics of brain disorders. *Nature Reviews Neuroscience* 16, 159-172.

Foulkes, L., Blakemore, S.J., 2018. Studying individual differences in human adolescent brain development. *Nature Neuroscience* 21, 315-323.

Fox, C.W., Roberts, S.J., 2012. A tutorial on variational Bayesian inference. *Artificial intelligence review* 38, 85-95.

Friston, K., Chu, C., Mourao-Miranda, J., Hulme, O., Rees, G., Penny, W., Ashburner, J., 2008. Bayesian decoding of brain images. *Neuroimage* 39, 181-205.

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. *Neuroimage* 34, 220-234.

Friston, K.J., Glaser, D.E., Henson, R.N.A., Kiebel, S., Phillips, C., Ashburner, J., 2002a. Classical and Bayesian inference in neuroimaging: Applications. *Neuroimage* 16, 484-512.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *Neuroimage* 19, 1273-1302.

Friston, K.J., Kahan, J., Biswal, B., Razi, A., 2014. A DCM for resting state fMRI. *Neuroimage* 94, 396-407.

Friston, K.J., Litvak, V., Oswal, A., Razi, A., Stephan, K.E., van Wijk, B.C.M., Ziegler, G., Zeidman, P., 2016. Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage* 128, 413-431.

Friston, K.J., Penny, W., 2003. Posterior probability maps and SPMs. *Neuroimage* 19, 1240-1249.

Friston, K.J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., Ashburner, J., 2002b. Classical and Bayesian inference in neuroimaging: theory. *Neuroimage* 16, 465-483.

Gal, Y., 2016. *Uncertainty in Deep Learning*. Department of Engineering. University of Cambridge, Cambridge.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D., 2013. *Bayesian Data Analysis*, 3 ed. CRC Press.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2014. *Bayesian data analysis*. Chapman & Hall/CRC Boca Raton, FL, USA.

Geschwind, D.H., Levitt, P., 2007. Autism spectrum disorders: developmental disconnection syndromes. *Current Opinion in Neurobiology* 17, 103-111.

Ghahramani, Z., 2015. Probabilistic machine learning and artificial intelligence. *Nature* 521, 452-459.

Goldman, S., 2013. Opinion: Sex, gender and the diagnosis of autism—A biosocial view of the male preponderance. *Research in Autism Spectrum Disorders* 7, 675-679.

Griffiths, T.L., Ghahramani, Z., 2011. The Indian Buffet Process: An Introduction and Review. *Journal of Machine Learning Research* 12, 1185-1224.

Groves, A.R., Beckmann, C.F., Smith, S.M., Woolrich, M.W., 2011. Linked independent component analysis for multimodal data fusion. *Neuroimage* 54, 2198-2217.

Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning*, 2nd ed. Springer, New York.

He, Y., Byrge, L., Kennedy, D.P., 2019. Non-replication of functional connectivity differences in ASD: a multi-site study. *bioRxiv*, 640797.

Hinne, M., Ambrogioni, L., Janssen, R.J., Heskes, T., van Gerven, M.A.J., 2014. Structurally-informed Bayesian functional connectivity analysis. *Neuroimage* 86, 294-305.

Huizinga, W., Poot, D., Vernooij, M., Rothschupkin, G., Ikram, M., Rueckert, D., Niessen, W., 2018. A spatio-temporal reference model of the aging brain. *NeuroImage* 169, 11-12.

Hull, L., Petrides, K., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., Mandy, W., 2017. “Putting on my best normal”: social camouflaging in adults with autism spectrum conditions. *J Autism Dev Disord* 47, 2519-2534.

Janssen, R.J., Jylanki, P., Kessels, R.P.C., van Gerven, M.A.J., 2015. Probabilistic model-based functional parcellation reveals a robust, fine-grained subdivision of the striatum. *Neuroimage* 119, 398-405.

Jaynes, E., 2003. *Probability Theory: The Logic of Science*. Cambridge University Press.

Jeffreys, H., 1946. AN INVARIANT FORM FOR THE PRIOR PROBABILITY IN ESTIMATION PROBLEMS. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* 186, 453-461.

Kanner, L., 1943. Autistic disturbances of affective contact. *Nervous child* 2, 217-250.

Kass, R.E., Steffey, D., 1989. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *Journal of the American Statistical Association* 84, 717-726.

Keary, C.J., Minshew, N.J., Bansal, R., Goradia, D., Fedorov, S., Keshavan, M.S., Hardan, A.Y., 2009. Corpus callosum volume and neurocognition in autism. *J Autism Dev Disord* 39, 834-841.

Kendall, A., Gal, Y., 2017. What uncertainties do we need in Bayesian deep learning for computer vision. *Neural Information Processing Systems*.

Kernbach, J., Satterthwaite, T., Bassett, D., Smallwood, J., Margulies, D., Krall, S., Shaw, P., Varoquaux, G., Thirion, B., Konrad, K., Bzdok, D., 2018. Shared Endo-phenotypes of Default Mode Dysfunction in Attention Deficit/Hyperactivity Disorder and Autism Spectrum Disorder. *Transl Psychiatry*.

Kim, J.E., Lyoo, I.K., Estes, A.M., Renshaw, P.F., Shaw, D.W., Friedman, S.D., Kim, D.J., Yoon, S.J., Hwang, J., Dager, S.R., 2010. Laterobasal amygdalar enlargement in 6- to 7-year-old children with autism spectrum disorder. *Arch Gen Psychiatry* 67, 1187-1197.

Kleinberg, J., 2002. An impossibility theorem for clustering. *NIPS* 15, 463-470.

Kruschke, J.K., 2011. *Doing Bayesian Data Analysis*. Elsevier, London, UK.

- Kruschke, J.K., Liddell, T.M., 2018. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon Bull Rev* 25, 178-206.
- Lai, M.-C., Lombardo, M.V., Auyeung, B., Chakrabarti, B., Baron-Cohen, S., 2015. Sex/gender differences and autism: setting the scene for future research. *Journal of the American Academy of Child & Adolescent Psychiatry* 54, 11-24.
- Lai, M.-C., Lombardo, M.V., Ruigrok, A.N., Chakrabarti, B., Auyeung, B., Szatmari, P., Happé, F., Baron-Cohen, S., Consortium, M.A., 2017a. Quantifying and exploring camouflaging in men and women with autism. *Autism* 21, 690-702.
- Lai, M.C., Lerch, J.P., Floris, D.L., Ruigrok, A.N., Pohl, A., Lombardo, M.V., Baron-Cohen, S., 2017b. Imaging sex/gender and autism in the brain: Etiological implications. *J Neurosci Res* 95, 380-397.
- Lima, C.H., Kwon, H.-H., Kim, J.-Y., 2016. A Bayesian beta distribution model for estimating rainfall IDF curves in a changing climate. *Journal of Hydrology* 540, 744-756.
- Liu, Y., Hayes, D.N., Nobel, A., Marron, J.S., 2008. Statistical Significance of Clustering for High-Dimension, Low-Sample Size Data. *Journal of the American Statistical Association* 103, 1281-1293.
- MacKay, D.J.C., 2003. Information theory, inference and learning algorithms. Cambridge university press.
- Mann, M.E., Lloyd, E.A., Oreskes, N., 2017. Assessing climate change impacts on extreme weather events: the case for an alternative (Bayesian) approach. *Climatic change* 144, 131-142.
- Marquand, A.F., Kia, S., Zabihi, M., Wolfers, T., Buitelaar, J.K., Beckmann, C.F., 2019. Conceptualizing mental disorders as deviations from normative functioning. *Molecular Psychiatry* in press.
- Marquand, A.F., Rezek, I., Buitelaar, J., Beckmann, C.F., 2016a. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biological Psychiatry* 80, 552-561.
- Marquand, A.F., Wolfers, T., Mennes, M., Buitelaar, J., Beckmann, C.F., 2016b. Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 1, 433-447.
- Matta, C.r.f., Massa, L., 2017. Notes on the energy equivalence of information. *The Journal of Physical Chemistry A* 121, 9131-9135.
- Miller, G.A., Chapman, J.P., 2001. Misunderstanding analysis of covariance. *Journal of Abnormal Psychology* 110, 40.
- Miller, K.L., Alfaro-Almagro, F., Bangarter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., 2016a. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci* 19, 1523.
- Miller, K.L., Alfaro-Almagro, F., Bangarter, N.K., Thomas, D.L., Yacoub, E., Xu, J.Q., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, J., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016b. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature Neuroscience* 19, 1523-1536.
- Mumford, J.A., Ramsey, J.D., 2014. Bayesian networks for fMRI: A primer. *Neuroimage* 86, 573-582.

Nacewicz, B.M., Dalton, K.M., Johnstone, T., Long, M.T., McAuliff, E.M., Oakes, T.R., Alexander, A.L., Davidson, R.J., 2006. Amygdala volume and nonverbal social impairment in adolescent and adult males with autism. *Arch Gen Psychiatry* 63, 1417-1428.

Neal, R.M., 1992. Bayesian mixture modeling. *workshop on maximum entropy and Bayesian methods of statistical analysis* Kluwer Academic Publishers.

Nikam, V.B., Meshram, B., 2013. Modeling rainfall prediction using data mining method: A Bayesian approach. 2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation. IEEE, pp. 132-136.

Nordahl, C.W., Simon, T.J., Zierhut, C., Solomon, M., Rogers, S.J., Amaral, D.G., 2008. Brief report: methods for acquiring structural MRI data in very young children with autism without the use of sedation. *J Autism Dev Disord* 38, 1581-1590.

Nosek, B.A., Aarts, A.A., Anderson, C.J., Anderson, J.E., Kappes, H.B., Collaboration, O.S., 2015. Estimating the reproducibility of psychological science. *Science* 349, aac4716-aac4716.

Pearl, J., Mackenzie, D., 2018. *The Book of Why: The New Science of Cause and Effect*. Basic Books.

Penny, W.D., Trujillo-Barreto, N.J., Friston, K.J., 2005. Bayesian fMRI time series analysis with spatial priors. *Neuroimage* 24, 350-362.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142-2154.

Rasmussen, C., 2000. *The Infinite Gaussian Mixture Model*. Neural Information Processing Systems.

Rasmussen, C.E., Williams, C., 2006. *Gaussian Processes for Machine Learning*. MIT Press.

Rosenberg, M.D., Finn, E.S., Scheinost, D., Papademetris, X., Shen, X.L., Constable, R.T., Chun, M.M., 2016. A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience* 19, 165-+.

Ruiz, F.J.R., Valera, I., Blanco, C., Perez-Cruz, F., 2014. Bayesian Nonparametric Comorbidity Analysis of Psychiatric Disorders. *Journal of Machine Learning Research* 15, 1215-1247.

Schaafsma, S.M., Pfaff, D.W., 2014. Etiologies underlying sex differences in autism spectrum disorders. *Frontiers in neuroendocrinology* 35, 255-271.

Schmidt, M.N., Winther, O., Hansen, L.K., 2009. *Bayesian Non-negative Matrix Factorization. Independent Component Analysis and Signal Separation*. Lecture Notes in Computer Science, Springer.

Scott, F.J., Baron-Cohen, S., Bolton, P., Brayne, C., 2002. Brief report prevalence of autism spectrum conditions in children aged 5-11 years in Cambridgeshire, UK. *Autism* 6, 231-237.

Seghier, M.L., Price, C., J, 2018. Interpreting and Utilising Intersubject Variability in Brain Function. *Trends in Cognitive Sciences* 22, 517-530.

Shalev-Shwartz, S., Ben-David, S., 2014. *Understanding machine learning: From theory to algorithms*. Cambridge University Press.

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E.J., Glasser, M.F., Uğurbil, K., Barch, D.M., Van Essen, D.C., Miller, K.L., 2015a. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience* 18, 1565-1567.

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E.J., Glasser, M.F., Uğurbil, K., Barch, D.M., Van Essen, D.C., Miller, K.L., 2015b. A positive-negative mode of population

covariation links brain connectivity, demographics and behavior. *Nat Neurosci* 18, 1565-1567.

Stephan, K.E., Schlagenhaut, F., Huys, Q.J.M., Raman, S., Aponte, E.A., Brodersen, K.H., Rigoux, L., Moran, R.J., Daunizeau, J., Dolan, R.J., 2017. Computational neuroimaging strategies for single patient predictions. *Neuroimage*.

Tang, S., Sun, N., Floris, D.L., Zhang, X., Di Martino, A., Yeo, B.T., 2019. Reconciling Dimensional and Categorical Models of Autism Heterogeneity: a Brain Connectomics & Behavioral Study. *bioRxiv*, 692772.

Tenenbaum, J.B., Kemp, C., Griffiths, T.L., Goodman, N.D., 2011. How to Grow a Mind: Statistics, Structure, and Abstraction. *Science* 331, 1279-1285.

Van Dijk, K.R., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431-438.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., Consortium, W.U.-M.H., 2013. The WU-Minn Human Connectome Project: An overview. *Neuroimage* 80, 62-79.

Volkow, N., Koob, G., Croyle, R., Bianchi, D., Gordon, J., Koroshetz, W., Pérez-Stable, E., Riley, W., Bloch, M., Conway, K., Deeds, B., Dowling, G., Grant, S., Howlett, K., Matochik, J., Morgan, G., Murray, M., Noronha, A., Spong, C., Wargo, E., Warren, K., Weiss, S., 2017. The conception of the ABCD study: From substance use to a broad NIH collaboration. *Developmental Cognitive Neuroscience* in press.

Wagenmakers, E.-J., Lee, M., Lodewyckx, T., Iverson, G.J., 2008. Bayesian versus frequentist inference. *Bayesian evaluation of informative hypotheses*. Springer, pp. 181-207.

Wolfers, T., Beckman, C.F., Hoogman, M., Buitelaar, J.K., Franke, B., Marquand, A., 2019a. Individual differences v. the average patient: mapping the heterogeneity in ADHD using normative models. *Psychological Medicine* In press.

Wolfers, T., Floris, D., Dinga, R., van Rooij, D., Isakoglou, C., Kia, S.M., Zabihi, M., Llera, A., Chowdanayaka, R., Kumar, V., Peng, H., Laid, C., Bataille, A., Dimitrova, R., Charman, T., Loth, E., Lai, M.-C., Jones, E., Baumeister, S., Moessnang, C., Banaschewski, T., Ecker, C., Dumas, G., O'Muircheartaigh, J., Murphy, D., Buitelaar, J.K., Marquand, A.F., Beckmann, C.F., 2019b. From pattern classification to stratification: towards conceptualizing the heterogeneity of Autism Spectrum Disorder. *Neuroscience and Biobehavioural Reviews* in press.

Woolrich, M.W., 2012. Bayesian inference in fMRI. *Neuroimage* 62, 801-810.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2004. Multilevel linear modelling for fMRI group analysis using Bayesian inference. *NeuroImage* 21, 1732-1747.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M., 2009. Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45, S173-S186.

Xia, C.H., Ma, Z.M., Ciric, R., Gu, S., Betzel, R.F., Kaczkurkin, A.N., Calkins, M.E., Cook, P.A., de la Garza, A.G., Vandekar, S.N., Cui, Z.X., Moore, T.M., Roalf, D.R., Ruparel, K., Wolf, D.H., Davatzikos, C., Gur, R.C., Gur, R.E., Shinohara, R.T., Bassett, D.S., Satterthwaite, T.D., 2018. Linked dimensions of psychopathology and connectivity in functional brain networks. *Nature Communications* 9.

Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., Fischl, B., 2014. Spurious group differences due to head motion in a diffusion MRI study. *Neuroimage* 88, 79-90.

Zabihi, M., Oldehinkel, M., Wolfers, T., Frouin, V., Goyard, D., Loth, E., Charman, T., Tilmann, J., Banaschewski, T., Dumas, G., Holt, R., Baron-Cohen, S., Durston, S., Bolte, S., Murphy, D.,

Ecker, C., Buitelaar, J.K., Beckmann, C.F., Marquand, A.F., 2018. Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* in press

Zhang, C., Butepage, J., Kjellstrom, H., Mandt, S., 2018. Advances in variational inference. *IEEE transactions on pattern analysis and machine intelligence*.