NeuroImage: Clinical 36 (2022) 103262

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Advancing brain network models to reconcile functional neuroimaging and clinical research

Xenia Kobeleva [a,b], Gaël Varoquaux [c], Alain Dagher [d], Mohit Adhikari [e], Christian Grefkes [f,g,h], Matthieu Gilson [i,j,k,*]

[a] Department of Neurology, University of Bonn, Bonn, Germany
[b] German Center for Neurodegenerative Diseases (DZNE) Bonn, Bonn, Germany
[c] INRIA Saclay, Paris, France
[d] Montreal Neurological Institute, McGill University, Montréal, Canada
[e] Bio-imaging Lab, University of Antwerp, Antwerp, Belgium
[f] Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[g] Institute of Neuroscience and Medicine (INM-1, INM-3), Research Centre Juelich, Juelich, Germany
[h] Department of Neurology, Goethe University Frankfurt, Frankfurt, Germany
[i] Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
[j] Center for Brain and Cognition, Department of Information and Telecommunication Technologies, Universitat Pompeu Fabra, Barcelona, Spain
[k] Institut de Neurosciences des Systèmes, Aix-Marseille University, Marseille, France

ARTICLE INFO

Keywords:
Whole-brain model
fMRI data
Diagnosis
Biomarker
Model interpretation
Neuropathologies

ABSTRACT

Functional magnetic resonance imaging (fMRI) captures information on brain function beyond the anatomical alterations that are traditionally visually examined by neuroradiologists. However, the fMRI signals are complex in addition to being noisy, so fMRI still faces limitations for clinical applications. Here we review methods that have been proposed as potential solutions so far, namely statistical, biophysical and decoding models, with their strengths and weaknesses. We especially evaluate the ability of these models to directly predict clinical variables from their parameters (predictability) and to extract clinically relevant information regarding biological mechanisms and relevant features for classification and prediction (interpretability). We then provide guidelines for useful applications and pitfalls of such fMRI-based models in a clinical research context, looking beyond the current state of the art. In particular, we argue that the clinical relevance of fMRI calls for a new generation of models for fMRI data, which combine the strengths of both biophysical and decoding models. This leads to reliable and biologically meaningful model parameters, which thus fulfills the need for simultaneous interpretability and predictability. In our view, this synergy is fundamental for the discovery of new pharmacological and interventional targets, as well as the use of models as biomarkers in neurology and psychiatry.

1. Neuroimaging in clinical practice and research

1.1. Evolution of neuroimaging diagnostic modalities in neuropsychiatric practice

Medical doctors originally studied and diagnosed diseases based solely on careful observation of symptoms and clinical examination. Several medical technologies, such as laboratory tests and radiographic techniques, had been initially met with skepticism given that they challenged the diagnostic authority of medical practitioners (Berger, 1999) and altered the patient-doctor relationship (Goold and Lipkin, 1999). History tends to repeat itself, and similar discussions are being held nowadays regarding the clinical application of machine learning to neuroimaging signals (Longoni et al., 2019).

Currently, diagnosis using magnetic resonance imaging (MRI) in neurology still mostly relies on qualitative analyses of brain structures and perfusion, which are restricted to anatomical alterations directly visible to the eye. With regard to psychiatry, clinical diagnoses predominantly rely on psychopathological explorations in combination with medical history and other factors as defined by specific diagnostic

* Corresponding author.
E-mail address: matthieu.gilson@univ-amu.fr (M. Gilson).

https://doi.org/10.1016/j.nicl.2022.103262
Received 3 August 2021; Received in revised form 26 October 2022; Accepted 6 November 2022
Available online 7 November 2022

manuals such as the Diagnostic and Statistical Manual of Mental Disorders (DSM), while neuroimaging measures remain at an exploratory stage in research. Thus, the clinical utilization of functional MRI (fMRI) does not stretch beyond the identification of localized functional areas for presurgical mapping before epilepsy or tumor surgery (Duffau, 2005; Silva et al., 2018). The broader acceptance of fMRI into clinical diagnostics is partly impeded by a general skepticism regarding the significance of the measured signals (Bennett et al., 2009; Kullmann, 2020), especially when considering the low signal-to-noise-ratio on a single subject level (Gorgolewski et al., 2013). While Blood Oxygen Level-Dependent (BOLD) signals have been clearly shown to indirectly reflect neuronal activity (Shmuel and Leopold, 2008), their relationship is complex and nonlinear. However, a good understanding of this complex relationship might be relevant for uncovering neuronal information processing and corresponding dysfunctions in neuropsychiatric diseases. In addition, BOLD signals are known to be contaminated by artifacts like head movement and respiration, which ideally requires careful preprocessing to access to the neuronal underlying contributions to fMRI signals (Fair et al. 2020; Lynch et al. 2021).

Despite their complex and noisy nature, fMRI data have shown potential for revealing brain dysfunction information not directly visible to the naked eye via an anatomical MRI scan. Various mathematical models have been developed for various goals such as reproducing the dynamics of fMRI signals and extracting information related to cognitive functions or pathological conditions. However, given that these models almost always involve highly multidimensional measures and/or parameterization, their interpretation for clinical application remains much more difficult compared to standard imaging diagnoses, in which normal anatomy is clearly established (e.g., identifying lesions or brain volumes). As such, whether more complex models would be sufficient to make fMRI clinically relevant, especially for neuropsychiatric disorders that still lack robust biological tests, still remains unknown.

This oriented review argues that better fMRI activity models are needed to enhance its clinical relevance. We start by providing guidelines on how to ask relevant questions for clinical neurology and psychiatry, applicable to any model in neurology. We then provide an overview of the model types based on or applied to functional brain networks, i.e., considering interactions between brain regions. In other words, we narrow our focus to network-oriented modeling approaches used to study neuropsychiatric diseases in a similar fashion as phenomenological functional connectivity (FC) (Bassett et al. 2008; Finn et al. 2014). We thus aim to go beyond earlier measures derived from empirical BOLD activity, such as voxel-level activity (mass univariate) or multivariate pattern analysis. Then, two distinct meanings of model *interpretability*—uncovering biological mechanisms versus identifying relevant features in a classifier—are discussed. In particular, we highlight the need for thorough validation of the relationship between fMRI-based functional connectivity (i.e., pairwise correlations of BOLD signals) and neuronal communication. We then discuss the *interpretability* of the models in contrast to *predictability*, herein understood as the ability of a model to predict clinical or cognitive variables from its parameters. Finally, we propose a promising approach for improving both *predictability* and *interpretability* that combines the strengths of the available models using parameters from biophysical models as features for decoding models (which will be described in more detail later). All in all, this review identifies roadblocks and advocates for more careful modeling and robust model validation.

1.2. Asking the right questions using translational research models

The traditional view of translational research mainly encompasses transferring innovative technologies from research into clinical practice ("from bench to bedside") and vice versa. Newer outlooks on translational research underline the importance of involving the community comprising patients, healthy populations, journalists, and medical practitioners (Cohrs et al., 2014; Forsythe et al., 2016). In the context of using data-driven models in translational research, collaboration between stakeholders can be substantially strengthened by asking precise and clinically relevant research questions. To this end, the "PICO" method (Richardson et al., 1995; Sackett et al., 2000) provides a framework comprising four components: P (population, patient, and/or problem), I (intervention, such as drugs, brain stimulation, diagnostic procedures, exposure, genetic factors), C (intervention for comparison), and O (outcome). A clinical research question is typically presented in the following format:

> *"In a given patient group, how does the intervention X differ from standard intervention Y in terms of outcome?"*
> *"In a given patient group, how does the diagnostic procedure X differ from standard diagnostic procedure Y in terms of outcome?"*

Subsequently, we argue that neuroimaging models could further enhance clinical research, with a collection of examples for each PICO category having been provided below.

1.2.1. Population, patient, and/or problem

One goal of fMRI-based diagnoses is to refine the categories of neuropsychiatric conditions (e.g., identify patient subgroups or individuals at risk) and compare them to standard categories based on other clinical measures. By doing so, neuroimaging models might help improve patient group selection for the comparison of different disease trajectories (Grefkes and Fink, 2014), namely identifying patients particularly suitable for specific (non-)pharmacological treatments according to treatment response (Dunlop and Mayberg, 2014) or predict different treatment outcomes. Such models might also help identify individuals at risk for certain diseases before the onset of clinical symptoms and thus enable trials on early preventive treatment, which are desperately needed for neurodegenerative diseases, such as Alzheimer's disease.

1.2.2. Intervention/comparison

fMRI-based measures or models can be tested against standard procedures (e.g., prognosis based on clinical variables) according to whether they improve prognostic or classification accuracy. Some applications of brain activity models have been shown to be clinically useful for assessing cortical functionality after stroke through transcranial magnetic stimulation combined with EEG (Tscherpel et al., 2020), studying regional cerebral diseases to enable targeted treatment in epilepsy surgery (Jirsa et al., 2017), or predicting potential complications of brain tumor resection (Woo et al., 2017). Furthermore, important research on the development of new treatments has aimed to uncover the neuronal mechanisms of neuropathologies, such as quantifying the local and global effects of medications (e.g., on neurotransmitters or neuronal excitability) on brain activity measured using fMRI. To address these concerns, next generation models require a systems medicine approach that possibly also involves animal models (Ren et al., 2014).

1.2.3. Outcome

Models might help create surrogate neuroimaging-based endpoints, such as better regeneration of damaged brain areas determined through MRI, which complement pure clinical endpoints (Prentice, 1989), e.g., long-term functional recovery. By identifying presymptomatic changes, these neuroimaging endpoints could reduce the duration of clinical trials and increase statistical power. To this end, models must capture longitudinal changes in patients' brain activity in order to characterize and determine disease outcomes.

2. Current and emerging models for extracting information from fMRI

Historically, research in modeling fMRI activity has followed two lines: one striving to reproduce and explain the spatiotemporal structure

of fMRI signals (e.g., at the whole-brain level) (statistical and biophysical models) and the other focused on extracting relevant information for the classification of patients and controls (decoding models). We now review these three broadly defined types of models in the context of functional brain networks. Their respective strengths with respect to clinical research objectives (as discussed in Section 1.2) pave the way for their combination (illustrated on Fig. 1 and discussed in Section 3).

2.1. Statistical generative models: reproducing the BOLD signal and its relevant patterns

Statistical models have long been used to evaluate the neuronal contribution to fMRI signals (Worsley et al., 1996, 1992). More recently, multivariate network models have been used to reproduce the structure of BOLD activity (Baldassarre et al., 2014; Bolton et al., 2018; Vidaurre et al., 2018). As illustrated in Fig. 1, the model parameters (e.g., network connectivity) are typically optimized to maximize the *goodness of fit* (see the vertical arrow), a measure of the match between empirical signals and their counterpart model signals. The model comes with hypotheses on the spatiotemporal structure of the observed data. Beyond goodness of fit, the model also adds value to clinical applications by determining how its estimated parameters can reliably reflect clinical conditions.

2.2. Biophysical models: uncovering biological mechanisms

Info box 1: Biophysical models

Biophysical models aim to formalize the link between brain anatomy (i.e., structural connectivity), neuronal dynamics and hemodynamics in the generation of BOLD signals. The modelling of hemodynamics to link neuronal activity to fMRI signals is a major difference with statistical models, which directly work at the level of BOLD signals. However, there is no strict boundary between biophysical and statistical models, with several studies falling somewhat in between both models. For instance, studies can use dynamic networks to directly generate the BOLD signals without

(*continued on next column*)

(*continued*)

hemodynamics while constraining their architecture with anatomical data (Gilson et al., 2016). A benefit of hemodynamics-based models, as with dynamic causal modeling, is that they can accommodate deviations from the standard canonical hemodynamic response function, which is especially useful in neurovascular disorders such as stroke (Grefkes et al., 2008).

With regard to model validation, many whole-brain biophysical models rely on functional connectivity rather than BOLD activation for model fitting (Deco et al., 2013; Ritter et al., 2013). While functional connectivity is typically interpreted as neuronal communication between brain regions, its relationship with neuronal activity has been much less validated than BOLD activation that has been thoroughly studied, e.g., in animal models (Bartels et al., 2008; David et al., 2008; Ekstrom, 2010). Efforts in that direction have been done to relate the propagation of BOLD signals with directional interactions in the early visual system (Gravel et al. 2020). We also note that recent animal studies have started to address the problem by comparing functional connectivity from fMRI and signals more directly related to neuronal activity like calcium imaging in mice (Chen et al. 2017; Lake et al. 2020). Likewise, relating functional connectivity to biological variables, such as synaptic efficacies, extends well beyond the general concept of communication between brain areas (Buckner, 2010).

Biophysical models, such as the dynamic causal model (Friston et al., 2003), the dynamic mean field model (Deco et al., 2013, Kobeleva et al., 2021a) or The Virtual Brain (Schirner et al. 2022), have been designed to elucidate the structure of observed BOLD activity, often formalizing the relationship of anatomical (or structural) connectivity with neural dynamics and/or relating neural dynamics to hemodynamics (Woolrich and Stephan, 2013, Glomb et al. 2021). Further details about their structure can be found in info box 1. In parallel, such models have been oriented towards explaining neuropathological BOLD activity by involving specific molecular mechanisms and biological variables, which aims to link with the fields of 'computational psychiatry' (Stephan and Mathys, 2014; Wang and Krystal, 2014) or 'computational neurology' (Maia and Frank, 2011).

As an example, recent biophysical models have designed the dynamics of neuronal populations mechanisms in a way that incorporate

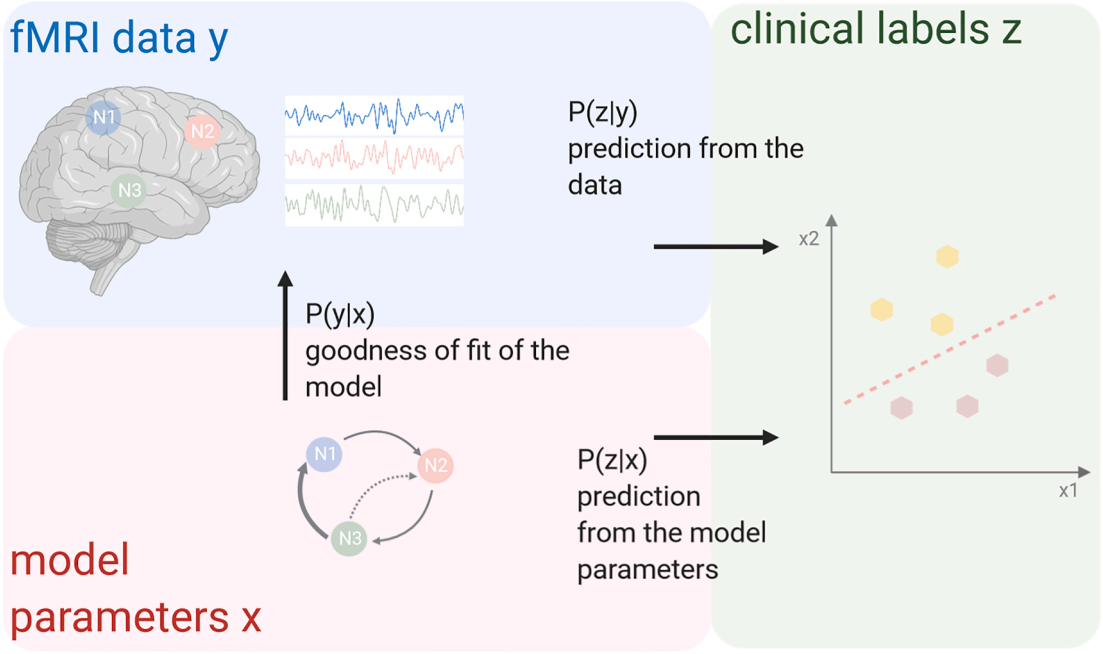

Fig. 1. Schematic representation of models applied to functional magnetic resonance imaging data. In a statistical or biophysical model (red box), parameters x are typically tuned to best reproduce the structure of Blood Oxygen Level-Dependent (BOLD) signals y (blue box), maximizing, for example, the likelihood P(y|x) of observing BOLD signals for parameterization x (see the vertical arrow). Predictive models aim to predict clinical labels z (green box; here, two categories are represented by different colored symbols) based on features y derived from BOLD signals (often a function applied to them, such as correlations for functional connectivity, instead of directly using the original signals themselves) or from the estimated model parameters x. Here P(z|…) refers to the probability that a new sample subject would belong to the category labeled z given the subject's features (y or x). In practice, this probability P(z|…) is estimated (and validated) using data with known labels, for example, by training a classifier. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

information from gene expression maps (Murray et al., 2018) or neurotransmitter concentration maps (Demirtaş et al., 2019). Their goal is to mechanistically describe the effect of local variables that can be measured from positron emission tomography or postmortem data on local and global brain activity measured using fMRI. This ability to trace back the effects of biological variables on fMRI measurements is an important part of *interpretability* as discussed in section 3.3.

Table 1 provides a comparison of different types of biophysical models with a focus on parameterization. We distinguish estimated parameters, which can be used as a signature of the brain activity state and are adjusted to fit the model to the observed fMRI signals, from other parameters that are derived directly from empirical data. Note that such models often involve mechanisms with additional free parameters, which may come from other studies in fundamental science or are heuristically determined. The models also differ in terms of number and types of estimated parameters, which determines the potential richness of the respective brain state signatures.

2.3. Decoding models: From fMRI images to clinical phenotypes

A short introduction about decoding models can be found in info box 2.

Info box 2: Decoding models

Decoding models have been increasingly used in machine learning to predict clinical or cognitive variables from high-dimensional fMRI data (Gao et al., 2019; Varoquaux et al., 2017). Practically, such models aim to map input *features*, which may include connectivity measures derived from BOLD signals or estimated parameters in a model as described earlier (see horizontal arrows in Fig. 1), to a *label* (e.g., a phenotype or diagnosis) or a *score* (e.g., cognitive test performance). The corresponding mathematical function from features to labels or scores is initially fit onto a *training set* of data. The generalizability of the prediction is then assessed using its performance on separate data (*test set*) in a process called *cross-validation* (Varoquaux et al., 2017).

In the same way that biophysical and statistical models are concerned with the goodness of fit of BOLD signals, decoding models are assessed with respect to their prediction performance, which we also refer to as *predictability*. A robust decoding model that predicts a phenotype should ideally provide information regarding which features contribute to the classification, another aspect of *interpretability* mentioned earlier (Pallarés et al., 2018; Yao et al., 2018). For instance, good motor recovery after a stroke can be classified according to primary motor cortex activity within the lesioned hemisphere measured using BOLD-fMRI on the first week after the stroke (Rehme et al., 2015).

Apart from using connectivity measures directly derived from BOLD data, region-wise or seed-based voxel-wise like functional connectivity (Naselaris et al., 2011; Richiardi et al., 2011), more complex metrics (e.g., time-varying functional connectivity) can be used as input for decoding models (Du et al., 2018). Decoding models can be applied to parameters estimated from statistical models, in which the model inversion can be seen as a preprocessing step (see the bottom horizontal arrow in Fig. 1). For instance, the Graphical Lasso can extract pairwise partial correlations among BOLD signals (Hoyos-Idrobo et al. 2017), resulting in a biomarker formed by a pair of regions with condition-specific changes in functional connectivity. More biologically-driven approaches can be designed by feeding machine learning tools with estimated parameters (e.g., effective connectivity) in a whole-brain dynamic model to predict subject individualities or cognitive states (Brodersen et al., 2011; Gilson et al., 2019; Pallarés et al., 2018). As a representation of BOLD signals, the estimated parameters may not only improve classification, but also yield a distinct interpretation (e.g., restricting the effective connections to anatomically connected regions).

3. Navigating between biophysical and decoding models: potentials and pitfalls

To address relevant questions in clinical research mentioned in Section 1 and establish clinically relevant and efficient biomarkers, we propose to bridge the gap between biophysical network and decoding models in order to combine their respective strengths, good interpretability together with good predictability. We now discuss the potentials and pitfalls of such unified modeling in the clinical environment, focusing on fMRI models.

Table 1

Building elements of biophysical models. Biophysical models involve parameters of different types: estimated by optimizing the model (from the fMRI data), directly derived from the data (other than fMRI), or free (e.g., from other experiments and not related to the fMRI data fitted by the model). Models can differ in terms of number and types of estimated parameters, ranging from one or two global parameters to hundreds or thousands of inter-regional connectivity estimates for models. Depending on combination of building elements, models can be used for different use-cases, e.g., linking structure and function in resting state, understanding task-evoked interaction of brain regions or, if additional neuroimaging information is included, combining neuroimaging modalities. An example use-case for modeling can be found in Fig. 2.

Use-case \ Parameter origin	Estimated	Data-derived	Free
Link structure and function (e.g., Deco et al. 2013)	Global coupling	Anatomical connectivity	Regional parameters (e.g., firing rate)
Combine different neuroimaging modalities (e.g., Murray et al., 2018; Demirtaş et al., 2019)	Global coupling	Anatomical connectivity Transmitter maps	Regional parameters (e.g., firing rate) Gene dynamics
Understand task-evoked activity (e.g., Gilson et al., 2019; Frässle et al., 2018)	Effective connectivity (per connection) Excitability (regional)		Global parameters (e.g., density of effective connectivity)

3.1. Robust decoding needs big data with reliable labels

Decoding models need accurate clinical labels, which can be challenging in itself. In neurology, a combination of medical history, physical examination, and laboratory tests has been used to provide standard classifications of symptoms and diseases that can be utilized as outcome parameters or labels (Bachmann et al., 2005). However, even gold standard clinical evaluations may have limited accuracy, which consequently limits the accuracy attainable by machine learning models. This concern is amplified in rare non-genetic neurological diseases. To start with, diagnosis is often not standardized (Haendel et al., 2020; Klimova et al., 2017), which affects the positive predictive value of a test (i.e., the proportion of true positive results within all positive test results) depending on the prevalence of a disease. Therefore, tests for rare diseases may show a relatively high absolute number of false positives despite excellent test sensitivity (Kohn et al., 2013; Lutgendorf and Stoll, 2016). Consequently, positive predictive values for rare diseases may drop drastically, subsequently compromising the labels for machine learning techniques.

In psychiatry, establishing reliable diagnostic labels is even more challenging considering that laboratory tests are often lacking and patients typically present within a wide range of disease severities. Thus, labeling is subject to the interpretation of clinicians who rely mostly on symptom-based diagnostic criteria (as proposed in disease classification manuals provided by different organizations, such as the Diagnostic and Statistical Manual of Mental Disorders, DSM).

However, the imperfection of clinical labels is not, by itself, an obstacle in developing useful neuroimaging biomarkers. Measures, such as repeated testing, can certainly improve the reliability of a training set given a sufficient sample size. Provided that such repeated tests do not have considerable variability, a classifier can learn to do better than the original clinical labels and thus enhance diagnostic accuracy (Wickenberg-Bolin et al., 2006).

Small sample sizes have been a well-known problem for machine learning techniques (Chu et al., 2012; Varoquaux, 2017). For complex data, machine learning techniques can only achieve good accuracy by using very large datasets to train the corresponding decoding model (Huf et al., 2014), similar to automated image recognition approaches using deep learning (Nguyen et al., 2020). Big data for neuropsychiatric diseases thus requires the alignment of diverse datasets in terms of neuroimaging (acquired using different scanners or protocols), as well as diagnostic criteria and clinical measures (Abraham et al., 2017; Karrer et al., 2019; Pomponio et al., 2020; Tax et al., 2019; Westeneng et al., 2018; Yamashita et al., 2019). Big-data approaches may not be easily applicable to rare or severe diseases, for which a sufficient number of cases is difficult to obtain. As such, a potential solution might be to group patients according to symptoms across different diagnostic categories or increase the number of measurements per subject to compensate for the small sample sizes (Krischer et al., 2014).

3.2. Personalized medicine requires tackling heterogeneity

A one-size-fits-all model of a disease might be helpful for understanding general disease pathophysiology and developing novel therapeutic targets at a group level. However, establishing a prognosis that can be useful across various patients requires a given model to integrate the evolution of different disease phenotypes and relevant patient characteristics. This means that the identifiability of the model parameters should be accessible in practice from data obtained from a single patient over, at most, a few fMRI recording sessions considering the well-known issues of signal-to-noise in fMRI data (Gorgolewski et al., 2013).

In a clinical context, fMRI studies easily suffer from selection bias. Most clinical trials have strict inclusion and exclusion criteria, resulting in patient groups that are not representative of the general patient population. In the field of stroke research, for instance, fMRI studies typically lack patients with very severe neurological deficits, given their inability to provide informed consent for a scientific study or lay still for a certain amount of time inside the scanner (Dani et al., 2008; Hotter et al., 2017; Kobeleva et al., 2021b). Therefore, many conclusions obtained from patient fMRI studies only apply to conscious, cooperative, mildly-to-moderately affected patients. Studies of neurodegenerative diseases face similar external-validity challenges: movement disorders some patients (e.g., tremors) interfere with MRI acquisition. Hence findings from fMRI clinical trials might be less applicable to patients with very severe neurological symptoms or atypical disease phenotypes (Mariani et al., 2019). Developing models on less severe symptoms that extrapolate to severe cases could be used to infer clinical outcomes for these (Salvalaggio et al., 2020).

Beyond the selection bias of fMRI studies, a given study cohort often displays a large heterogeneity. Numerous patients suffer from multiple diseases (e.g., a patient with dementia having recurrent strokes, a pacemaker, and lung cancer or a patient with depression, as well as personality and addiction disorders). The problem with non-homogeneous samples is further complicated by inter-individual differences in drug treatment, with some drugs interfering with BOLD signals such as certain neuroleptics (Röder et al., 2010), as well as the complex interactions resulting from taking multiple drugs. Such multimorbidity and polypharmacy has remained a considerable problem in clinical studies, that can only be addressed through very large samples, which make studies exceedingly expensive, long-lasting, and complicated to organize. One important step to address such concerns has been population-based approaches like the UK Biobank study (Elliott et al., 2018; Ward et al., 2019) and the use of normative models (Marquand et al., 2016). However, remains the problem of the considerably large fraction of patients that cannot be included into an MRI study due to contraindications, such as non-MRI-compatible pacemakers, metal splints, or implanted electrodes for deep brain stimulations. For such patients, other techniques for obtaining brain activity should be explored, such as near infrared spectroscopy or EEG-based techniques. Different neuroimaging signals must then be aligned, for instance using functional connectivity from fMRI and EEG in the same analysis.

Lastly, the fMRI signal can be altered in conditions with neurovascular alterations as encountered in the case of atherosclerosis, leading to narrowing or occlusion of vessels. These conditions can be frequently encountered in older populations and may be clinically silent and cannot be inferred from BOLD-sensitive images (e.g., gradient-echo T2*) post-hoc.

Therefore, fMRI studies with patients or older individuals should ideally always include diagnostic scans to ensure the validity of BOLD signals computed from fMRI volumes. Currently, using anatomical information (e.g., atrophy) plotted onto a normative connectome to predict functional connectivity patterns can be used as a workaround (Corp et al., 2019; Horn et al., 2017). Ideally, future models should incorporate corrected function of the hemodynamic response in the case of lesions or atrophy.

3.3. Bridging interpretability and predictability to uncover potential therapies for pathological neuronal mechanisms

Beyond giving accurate prediction, models for clinical research should also be interpretable (see Fig. 2). We discuss two distinct meanings of interpretability: understanding the important features and unraveling biological mechanisms. The first meaning of interpretability can be achieved by understanding informative features of decoding models and leads to the identification of relevant brain regions for disease understanding (i.e., those that show abnormal brain function) (Abraham et al., 2017; Hoyos-Idrobo et al., 2018). For instance, interpretable architectures can reconstruct inverse images of interesting labels from deep neural networks into brain anatomical regions (Böhle et al., 2019). This type of interpretability has several benefits in the context of translational research. It can be used to detect hidden biases

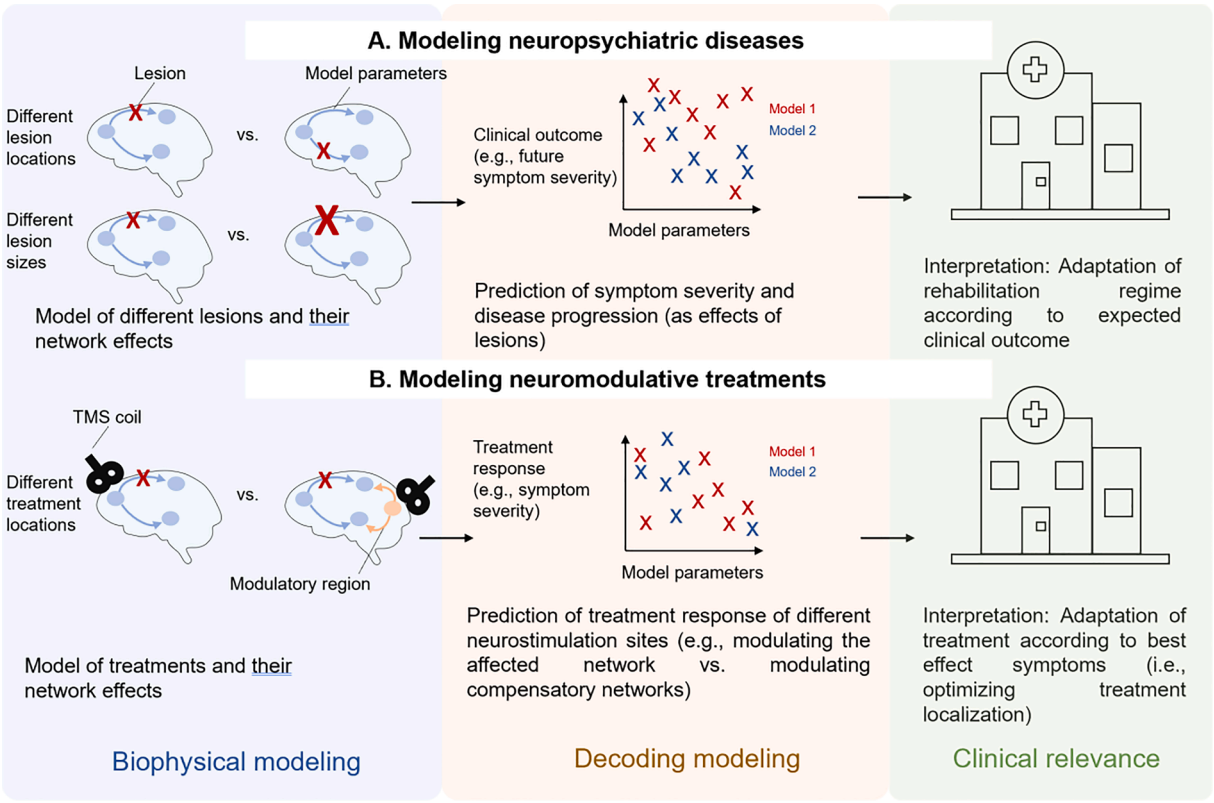

Fig. 2. Combining prediction and interpretation by using model parameters from biophysical models as input parameters of decoding models. A. As an example use-case we consider a biophysical model of an altered network interactions due to a stroke (*blue column*). Here, different lesion locations or sizes may have differential effects on network dynamics (in terms of altered model parameters; note that here only one dimension is represented). The resulting parameter signature (as a representation of BOLD signals) can be used as input features for a decoding model (*orange column*) to predict the clinical outcome after a stroke (*green column*). This also enables a clinically meaningful model comparison, for instance to assess whether a particular biophysical model improves interpretability while retaining sufficient predictability as compared to more phenomenological approaches like FC analysis, in order to achieve the desired sweet spot between interpretability and predictability. B. Besides interpreting and predicting the effect of lesions, biophysical models can further be used to evaluate the effects of different treatments, e.g., TMS of a directly affected region versus a third-party modulatory brain region (*blue column*). In both cases, as before, the effect of the different treatments can be quantified by feeding the parameters of the biophysical model as features for the decoding model (*orange column*), for instance to identify the reorganization of interactions in subnetworks in relation to clinical variables. This pipeline thus potentially allows for a comparison between different mechanistic model-based interpretations of treatments and their effects on patients' symptoms (*green column*). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

through careful model validation by applying the model to new data (e.g., the influence of other pathophysiological processes). Furthermore, the evaluation and correction of potential prediction errors allows clinicians to consider disease severity and adverse effects when choosing the appropriate treatment (e.g., delaying treatment due to the risk of developing a brain tumor vs undergoing a risky treatment for migraine headache).

The second meaning of interpretability, in the sense of mechanistic interpretation, can be achieved by integrating biologically meaningful parameters obtained from other imaging modalities (e.g., neurotransmitter concentration) into the modelled brain dynamics. This is key to developing pharmaceutics based on the manipulation of physiological mechanisms that lead to the altered brain dynamics. However, to be used in clinical research, models including different data sources for parameter estimation should be aligned. As an example, effective connectivity may quantitatively and qualitatively differ according to whether or not the model includes neurotransmitter concentrations. In biophysical whole-brain models, the structure of empirical BOLD signals (e.g., functional connectivity) can be determined via the estimated parameters whose numbers may vary dramatically across different models. For instance, effective connectivity in a whole-brain model can provide a rich representation of BOLD dynamics involving thousands of estimated weights (Frässle et al., 2018; Gilson et al., 2019), similar to the statistical models mentioned previously. In contrast, model fitting may involve only a few parameters, such as a single global coupling parameter in the original mean field model (Deco et al., 2013), or several parameters that determine highly nonlinear nodal dynamics (Proix et al., 2017; Pohl et al. 2022).

In any case, such biophysical models have to be identifiable with respect to their fitting procedure (Wilson and Collins, 2019). This should be checked using parameter recovery (e.g., creating simulated data with known parameters and fitting a model to estimate these parameters), to ensure that the estimated parameters in the model unambiguously represent the reproduced BOLD structure. These estimated parameters can be used as features in decoding models, which establishes solid interpretations on brain activity. *Identifiability* is all the more crucial when parameters are estimated in increasingly complex models, such as those that aim to comprehensively describe pathological alterations of neuronal activity involving neurotransmitters (see Table 1). Free parameters with no relationship to the data are often present when linking diverse data types in a model, which may affect the estimation of other parameters and necessitates control using robustness checks. These checks ensure similar results across different variations of biophysical models and generalizability of estimated parameters using out-of-sample data.

Given robust parameters across different data types, biophysical models could be synergistically combined with decoding models. In that way their respective strengths can be combined in a next generation of

biophysical models that inform on the biological mechanisms which are relevant for prediction, as illustrated in Fig. 2. This will allow for a translation from pure machine-learning diagnosis to a further level of model-based prediction. In our view this new generation of models will be able to achieve a balance between interpretability (both in terms of identification of relevant features and mechanistic interpretation) and predictability of clinical outcome. The key is to identify biologically meaningful data-based and estimated parameters that represent the brain activity state. Using these estimated parameters as features for decoding models can build dynamical signatures. A step in that direction has been made in the field by designing models of effective connectivity that give good prediction accuracy (Brodersen et al., 2011; Frässle et al., 2018; Gilson et al., 2019). As an example, the effective connectivity in an anatomically constrained model can quantify the modulation of white-matter pathways in a condition-specific fashion (related to a neuropathology or cognitive task). This improves the interpretation of the model parameters as compared to FC that describes the statistical relationship for all possible pairs of brain regions, irrespective of whether they are anatomically connected or not. It can also enhance the predictability as measured by the decoding accuracy, by reducing the number of estimated parameters (i.e., by taking into account anatomical pathways when estimating connections between regions).

In the illustrative example for stroke in Fig. 2A, patients exhibit lesions at different locations and of various sizes. Prognosis consists in predicting the future evolution of the patient, e.g., his/her recovery in terms of cognitive deficits after a year from fMRI signals acquired during the acute phase (a few weeks after the stroke). The goal of the biophysical model is to infer parameters that characterize the neuronal dynamics and interactions from the empirical fMRI, as recently demonstrated using effective connectivity for cortico-cortical communication (Adhikari et al. 2021). Going a step further, the potential effect of different treatments could be evaluated using a biophysical model to identify treatment-induced changes in cortical subnetworks from the estimated parameters, and using subsequently these parameters in a decoding model to test the effect on the patient's recovery or disease progression (i.e., expected symptom severity in the future). In both cases, the combined modeling makes a triple link between the perturbation (stroke lesion, TMS treatment), the cortical network dynamics and the patient's condition.

Using these new types of models will allow exploring meaningful therapeutic pharmacological or interventional targets beyond the localization of brain regions with strongest changes in BOLD activity or functional connectivity interactions. Of course, increasing the complexity of biophysical models to incorporate e.g., many mechanisms like excitation-inhibition balance and neuromodulation leads to highly non-linear network dynamics and the estimation of many parameters is notoriously difficult in this case, especially for connectivity parameters. Therefore, the challenge lies in carefully designing the model such that it remains tractable for parameter estimation (necessary for predictability) while combining biologically relevant mechanisms (for interpretability). Importantly, to ensure superiority of model-based prediction against other predictive pipelines, predictability of a model should be tested against a gold standard (if it exists) or against/ in combination with other neuroimaging modalities (e.g., atrophy or structural measures- which are easier to obtain than model-related features).

4. Perspective: Bringing models into clinical practice

Looking to the future, large neuroimaging datasets covering many diseases will bring better models of fMRI activity, that contribute to improving diagnosis and treatment in neurology and psychiatry. Models suited to clinical applications must not only capture better BOLD activity, for instance via functional connectivity, but also the influence of other biological factors, such as neurotransmitters, proteins, and genes. A critical limitation for clinical applicability lies in the noisy and heterogeneous nature of patients' data. As models improve to capture this heterogeneity, they will gradually become useful tools for clinicians, supporting decision making, although they should not replace clinical reasoning. They could constitute an additional element within a set of established diagnostic procedures and might increase the certainty of a diagnostic classification or risk assessment through the use of multimodal decision support systems. To improve functionality, these models should be connected with more complete electronic health records to extract additional patient data and visually present the results of the analysis, for example, through secure and flexible web services (Khalilia et al., 2015). Broader involvement would require additional training for dedicated medical personnel in model-based data analyses, in the same manner as MRI physicists.

Using increasingly sensitive models of brain activity for clinical practice has opportunities for early treatment but also ethical consequences. Similar to the situation of presymptomatic genetic testing, doctors and patients might face the challenge of discussing a diagnosis of a presymptomatic neuropsychiatric disease based solely on a possibly complex model without any visible brain pathology. What's more, there are direct consequences on the daily life of patients affecting career choice, family planning, and health insurance policies, as well as have psychological consequences (Godino et al., 2016; Tibben et al., 1997). Therefore, ethical guidelines must be developed alongside the models to assess the potential benefit and risks of communicating the model results to the patients.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was supported by an add-on fellowship of the Joachim Herz Foundation of XK. CG is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 431549029 – SFB 1451.

References

Abraham, A., Milham, M.P., Di Martino, A., Craddock, R.C., Samaras, D., Thirion, B., Varoquaux, G., 2017. Deriving reproducible biomarkers from multi-site resting-state data: an Autism-based example. NeuroImage 147, 736–745. https://doi.org/10.1016/j.neuroimage.2016.10.045.

Adhikari, M.H., Griffis, J., Siegel, J.S., Thiebaut de Schotten, M., Deco, G., Instabato, A., Gilson, M., Corbetta, M., 2021. Effective connectivity extracts clinically relevant prognostic information from resting state activity in stroke. Brain Commun. 3 (fcab233).

Baldassarre, A., Ramsey, L., Hacker, C.L., Callejas, A., Astafiev, S.V., Metcalf, N.V., Zinn, K., Rengachary, J., Snyder, A.Z., Carter, A.R., Shulman, G.L., Corbetta, M., 2014. Large-scale changes in network interactions as a physiological signature of spatial neglect. Brain 137, 3267–3283. https://doi.org/10.1093/brain/awu297.

Bartels, A., Logothetis, N.K., Moutoussis, K., 2008. fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci. 31, 444–453. https://doi.org/10.1016/j.tins.2008.06.004.

Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia. J. Neurosci. 28, 9239–9248.

Bennett, C.M., Wolford, G.L., Miller, M.B., 2009. The principled control of false positives in neuroimaging. Soc. Cogn. Affect. Neurosci. 4, 417–422. https://doi.org/10.1093/scan/nsp053.

Berger, D., 1999. A brief history of medical diagnosis and the birth of the clinical laboratory. Part 1–Ancient times through the 19th century. MLO. Med. Lab. Obs. 31, 28–30, 32, 34–40.

Böhle, M., Eitel, F., Weygandt, M., Ritter, K., 2019. Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification. Front. Aging Neurosci. 11, 194. https://doi.org/10.3389/fnagi.2019.00194.

Bolton, T.A.W., Tarun, A., Sterpenich, V., Schwartz, S., Van De Ville, D., 2018. Interactions between large-scale functional brain networks are captured by sparse coupled HMMs. IEEE Trans. Med. Imaging 37, 230–240. https://doi.org/10.1109/TMI.2017.2755369.

Brodersen, K.H., Schofield, T.M., Leff, A.P., Ong, C.S., Lomakina, E.I., Buhmann, J.M., Stephan, K.E., 2011. Generative embedding for model-based classification of fMRI data. PLoS Comput. Biol. 7, e1002079.

Buckner, R.L., 2010. Human functional connectivity: new tools, unresolved questions. Proc. Natl. Acad. Sci. 107, 10769. https://doi.org/10.1073/pnas.1005987107.

Chen, L.M., Yang, P.-F., Wang, F., Mishra, A., Shi, Z., Wu, R., Wu, T.-L., Wilson, G.H., Ding, Z., Gore, J.C., 2017. Biophysical and neural basis of resting state functional connectivity: evidence from non-human primates. Magn. Reson. Imaging 39, 71–81.

Chu, C., Hsu, A.-L., Chou, K.-H., Bandettini, P., Lin, C., 2012. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. NeuroImage 60, 59–70. https://doi.org/10.1016/j.neuroimage.2011.11.066.

Cohrs, R.J., Martin, T., Ghahramani, P., Bidaut, L., Higgins, P.J., Shahzad, A., 2014. Translational medicine definition by the european society for translational medicine. Eur. J. Mol. Clin. Med. 2, 86. https://doi.org/10.1016/j.nhtm.2014.12.002.

Corp, D.T., Joutsa, J., Darby, R.R., Delnooz, C.C.S., van de Warrenburg, B.P.C., Cooke, D., Prudente, C.N., Ren, J., Reich, M.M., Batla, A., Bhatia, K.P., Jinnah, H.A., Liu, H., Fox, M.D., 2019. Network localization of cervical dystonia based on causal brain lesions. Brain J. Neurol. 142, 1660–1674. https://doi.org/10.1093/brain/awz112.

Dani, K.A., McCormick, M.T., Muir, K.W., 2008. Brain lesion volume and capacity for consent in stroke trials: potential regulatory barriers to the use of surrogate markers. Stroke 39, 2336–2340. https://doi.org/10.1161/STROKEAHA.107.507111.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., Depaulis, A., 2008. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biol. 6 https://doi.org/10.1371/journal.pbio.0060315.

Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G.L., Hagmann, P., Corbetta, M., 2013. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. J. Neurosci. Off. J. Soc. Neurosci. 33, 11239–11252. https://doi.org/10.1523/JNEUROSCI.1091-13.2013.

Demirtaş, M., Burt, J.B., Helmer, M., Ji, J.L., Adkinson, B.D., Glasser, M.F., Van Essen, D. C., Sotiropoulos, S.N., Anticevic, A., Murray, J.D., 2019. Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. Neuron 101, 1181–1194. e13. https://doi.org/10.1016/j.neuron.2019.01.017.

Du, Y., Fu, Z., Calhoun, V.D., 2018. Classification and prediction of brain disorders using functional connectivity: promising but challenging. Front. Neurosci. 12 https://doi.org/10.3389/fnins.2018.00525.

Duffau, H., 2005. Lessons from brain mapping in surgery for low-grade glioma: insights into associations between tumour and brain plasticity. Lancet Neurol. 4, 476–486. https://doi.org/10.1016/S1474-4422(05)70140-X.

Dunlop, B.W., Mayberg, H.S., 2014. Neuroimaging-based biomarkers for treatment selection in major depressive disorder. Dialogues Clin. Neurosci. 16, 479–490.

Ekstrom, A., 2010. How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. Brain Res. Rev. 62, 233–244. https://doi.org/10.1016/j.brainresrev.2009.12.004.

Elliott, L.T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K.L., Douaud, G., Marchini, J., Smith, S.M., 2018. Genome-wide association studies of brain imaging phenotypes in UK Biobank. Nature 562, 210. https://doi.org/10.1038/s41586-018-0571-7.

Fair, D.A., Miranda-Dominguez, O., Snyder, A.Z., Perrone, A., Earl, E.A., Van, A.N., Koller, J.M., Feczko, E., Tisdall, M.D., van der Kouwe, A., Klein, R.L., Mirro, A.E., Hampton, J.M., Adeyemo, B., Laumann, T.O., Gratton, C., Greene, D.J., Schlaggar, B. L., Hagler, D.J., Watts, R., Garavan, H., Barch, D.M., Nigg, J.T., Petersen, S.E., Dale, A.M., Feldstein-Ewing, S.W., Nagel, B.J., Dosenbach, N.U.F., 2020. Correction of respiratory artifacts in MRI head motion estimates. NeuroImage 208, 116400.

Finn, E.S., Shen, X., Holahan, J.M., Scheinost, D., Lacadie, C., Papademetris, X., Shaywitz, S.E., Shaywitz, B.A., Constable, R.T., 2014. Disruption of functional networks in dyslexia: a whole-brain, data-driven analysis of connectivity. Biological Psychiatry 76, 397–404.

Forsythe, L.P., Ellis, L.E., Edmundson, L., Sabharwal, R., Rein, A., Konopka, K., Frank, L., 2016. Patient and stakeholder engagement in the PCORI pilot projects: description and lessons learned. J. Gen. Intern. Med. 31, 13–21. https://doi.org/10.1007/s11606-015-3450-z.

Frässle, S., Lomakina, E.I., Kasper, L., Manjaly, Z.M., Leff, A., Pruessmann, K.P., Buhmann, J.M., Stephan, K.E., 2018. A generative model of whole-brain effective connectivity. NeuroImage 179, 505–529. https://doi.org/10.1016/j.neuroimage.2018.05.058.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19, 1273–1302.

Gao, S., Greene, A.S., Constable, R.T., Scheinost, D., 2019. Combining multiple connectomes improves predictive modeling of phenotypic measures. NeuroImage 201, 116038. https://doi.org/10.1016/j.neuroimage.2019.116038.

Gilson, M., Moreno-Bote, R., Ponce-Alvarez, A., Ritter, P., Deco, G., 2016. Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome. PLOS Comput. Biol. 12, e1004762.

Gilson, M., Zamora-López, G., Pallarés, V., Adhikari, M.H., Senden, M., Campo, A.T., Mantini, D., Corbetta, M., Deco, G., Insabato, A., 2019. Model-based whole-brain effective connectivity to study distributed cognition in health and disease. Netw. Neurosci. 4, 338–373. https://doi.org/10.1162/netn_a_00117.

Glomb, K., Kringelbach, M.L., Deco, G., Hagmann, P., Pearson, J., Atasoy, S., 2021. Functional harmonics reveal multi-dimensional basis functions underlying cortical organization. Cell Reports. 36, 109554.

Godino, L., Turchetti, D., Jackson, L., Hennessy, C., Skirton, H., 2016. Impact of presymptomatic genetic testing on young adults: a systematic review. Eur. J. Hum. Genet. 24, 496–503. https://doi.org/10.1038/ejhg.2015.153.

Goold, S.D., Lipkin, M., 1999. The doctor-patient relationship. J. Gen. Intern. Med. 14, S26–S33. https://doi.org/10.1046/j.1525-1497.1999.00267.x.

Gorgolewski, K.J., Storkey, A.J., Bastin, M.E., Whittle, I., Pernet, C., 2013. Single subject fMRI test–retest reliability metrics and confounding factors. NeuroImage 69, 231–243. https://doi.org/10.1016/j.neuroimage.2012.10.085.

Gravel, N., Renken, R.J., Harvey, B.M., Deco, G., Cornelissen, F.W., Gilson, M., 2020. Propagation of BOLD activity reveals task-dependent directed interactions across human visual cortex. Cerebral Cortex. 30, 5899–5914.

Grefkes, C., Fink, G.R., 2014. Connectivity-based approaches in stroke and recovery of function. Lancet Neurol. 13, 206–216. https://doi.org/10.1016/S1474-4422(13)70264-3.

Grefkes, C., Nowak, D.A., Eickhoff, S.B., Dafotakis, M., Küst, J., Karbe, H., Fink, G.R., 2008. Cortical connectivity after subcortical stroke assessed with functional magnetic resonance imaging. Ann. Neurol. 63, 236–246. https://doi.org/10.1002/ana.21228.

Haendel, M., Vasilevsky, N., Unni, D., Bologa, C., Harris, N., Rehm, H., Hamosh, A., Baynam, G., Groza, T., McMurry, J., Dawkins, H., Rath, A., Thaxon, C., Bocci, G., Joachimiak, M.P., Köhler, S., Robinson, P.N., Mungall, C., Oprea, T.I., 2020. How many rare diseases are there? Nat. Rev. Drug Discov. 19, 77–78. https://doi.org/10.1038/d41573-019-00180-y.

Horn, A., Reich, M., Vorwerk, J., Li, N., Wenzel, G., Fang, Q., Schmitz-Hübsch, T., Nickl, R., Kupsch, A., Volkmann, J., Kühn, A.A., Fox, M.D., 2017. Connectivity predicts deep brain stimulation outcome in Parkinson disease. Ann. Neurol. 82, 67–78. https://doi.org/10.1002/ana.24974.

Hotter, B., Ulm, L., Hoffmann, S., Katan, M., Montaner, J., Bustamante, A., Meisel, A., 2017. Selection bias in clinical stroke trials depending on ability to consent. BMC Neurol. 17 https://doi.org/10.1186/s12883-017-0989-9.

Hoyos-Idrobo, A., Varoquaux, G., Schwartz, Y., Thirion, B., 2018. FReM – Scalable and stable decoding with fast regularized ensemble of models. NeuroImage, New advances in encoding and decoding of brain signals 180, 160–172. https://doi.org/10.1016/j.neuroimage.2017.10.005.

Huf, W., Kalcher, K., Boubela, R.N., Rath, G., Vecsei, A., Filzmoser, P., Moser, E., 2014. On the generalizability of resting-state fMRI machine learning classifiers. Front. Hum. Neurosci. 8 https://doi.org/10.3389/fnhum.2014.00502.

Jirsa, V.K., Proix, T., Perdikis, D., Woodman, M.M., Wang, H., Gonzalez-Martinez, J., Bernard, C., Bénar, C., Guye, M., Chauvel, P., Bartolomei, F., 2017. The Virtual Epileptic Patient: Individualized whole-brain models of epilepsy spread. NeuroImage 145, 377–388. https://doi.org/10.1016/j.neuroimage.2016.04.049.

Karrer, T.M., Bassett, D.S., Derntl, B., Gruber, O., Aleman, A., Jardri, R., Laird, A.R., Fox, P.T., Eickhoff, S.B., Grisel, O., Varoquaux, G., Thirion, B., Bzdok, D., 2019. Brain-based ranking of cognitive domains to predict schizophrenia. Hum. Brain Mapp. 40, 4487–4507. https://doi.org/10.1002/hbm.24716.

Khalilia, M., Choi, M., Henderson, A., Iyengar, S., Braunstein, M., Sun, J., 2015. Clinical predictive modeling development and deployment through FHIR web services. AMIA. Annu. Symp. Proc. 2015, 717–726.

Klimova, B., Storek, M., Kuca *, M.V. and K., 2017. Global View on Rare Diseases: A Mini Review [WWW Document]. Curr. Med. Chem.

Kobeleva, X., López-González, A., Kringelbach, M.L., Deco, G., 2021a. Revealing the relevant spatiotemporal scale underlying whole-brain dynamics. Front Neurosci. 15, 715861.

Kobeleva X, Machts J, Veit M, Vielhaber S, Petri S, Schoenfeld MA. 2021b. Brain activity is contingent on neuropsychological function in a functional magnetic resonance imaging study of verbal working memory in amyotrophic lateral sclerosis. Eur J Neurol. ene.14957.

Kohn, M.A., Carpenter, C.R., Newman, T.B., 2013. Understanding the direction of bias in studies of diagnostic test accuracy. Acad. Emerg. Med. 20, 1194–1206. https://doi.org/10.1111/acem.12255.

J. Gen. Intern. Med. 29 Suppl 3, S739-744. doi: 10.1007/s11606-014-2894-x.

Kullmann, D.M., 2020. Editorial. Brain 143.

Lake, E.M.R., Ge, X., Shen, X., Herman, P., Hyder, F., Cardin, J.A., Higley, M.J., Scheinost, D., Papademetris, X., Crair, M.C., Constable, R.T., 2020. Simultaneous cortex-wide fluorescence Ca2+ imaging and whole-brain fMRI. Nat. Methods 17, 1262–1271.

Longoni, C., Bonezzi, A., Morewedge, C.K., 2019. Resistance to medical artificial intelligence. J. Consum. Res. 46, 629–650. https://doi.org/10.1093/jcr/ucz013.

Lutgendorf, M.A., Stoll, K.A., 2016. Why 99% may not be as good as you think it is: limitations of screening for rare diseases. J. Matern.-Fetal Neonatal Med. Off. J. Eur. Assoc. Perinat. Med. Fed. Asia Ocean. Perinat. Soc. Int. Soc. Perinat. Obstet. 29, 1187–1189. doi: 10.3109/14767058.2015.1039977.

Lynch, C.J., Voss, H.U., Silver, B.M., Power, J.D., 2021. On measuring head motion and effects of head molds during fMRI. NeuroImage 225, 117494.

Maia, T.V., Frank, M.J., 2011. From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. 14, 154–162. https://doi.org/10.1038/nn.2723.

Mariani, L.-L., Guimarães-Costa, R., Grabli, D., Le Toullec, B., Cormier-Dequaire, F., Degos, B., Dubois, B., Vidailhet, M., Lacomblez, L., Corvol, J.-C., 2019. Are PSP patients included in clinical trials representative of the general PSP population? Parkinsonism Relat. Disord. 66, 202–206. https://doi.org/10.1016/j.parkreldis.2019.07.012.

Marquand, A.F., Rezek, I., Buitelaar, J., Beckmann, C.F., 2016. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. Biol. Psychiatry 80, 552–561. https://doi.org/10.1016/j.biopsych.2015.12.023.

Murray, J.D., Demirtaş, M., Anticevic, A., 2018. Biophysical modeling of large-scale brain dynamics and applications for computational psychiatry. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 3, 777–787. https://doi.org/10.1016/j.bpsc.2018.07.004.

Naselaris, T., Kay, K.N., Nishimoto, S., Gallant, J.L., 2011. Encoding and decoding in fMRI. NeuroImage 56, 400–410. https://doi.org/10.1016/j.neuroimage.2010.07.073.

Nguyen, M., He, T., An, L., Alexander, D.C., Feng, J., Yeo, B.T.T., 2020. Predicting Alzheimer's disease progression using deep recurrent neural networks. NeuroImage 222, 117203. https://doi.org/10.1016/j.neuroimage.2020.117203.

Pallarés, V., Insabato, A., Sanjuán, A., Kühn, S., Mantini, D., Deco, G., Gilson, M., 2018. Extracting orthogonal subject- and condition-specific signatures from fMRI data using whole-brain effective connectivity. NeuroImage 178, 238–254. https://doi.org/10.1016/j.neuroimage.2018.04.070.

Pohl, E.D.R., Upadhyay, N., Kobeleva, X., Purrer, V., Maurer, A., Keil, V.C., Kindler, C., Borger, V., Pieper, C.C., Groetz, S., Scheef, L., Maciaczyk, J., Schild, H., Vatter, H., Klockgether, T., Radbruch, A., Attenberger, U., Wüllner, U., Boecker, H., 2022. Coherent structural and functional network changes after thalamic lesions in essential tremor. Mov. Disord.

Pomponio, R., Erus, G., Habes, M., Doshi, J., Srinivasan, D., Mamourian, E., Bashyam, V., Nasrallah, I.M., Satterthwaite, T.D., Fan, Y., Launer, L.J., Masters, C.L., Maruff, P., Zhuo, C., Völzke, H., Johnson, S.C., Fripp, J., Koutsouleris, N., Wolf, D.H., Gur, R., Gur, R., Morris, J., Albert, M.S., Grabe, H.J., Resnick, S.M., Bryan, R.N., Wolk, D.A., Shinohara, R.T., Shou, H., Davatzikos, C., 2020. Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. NeuroImage 208, 116450. https://doi.org/10.1016/j.neuroimage.2019.116450.

Prentice, R.L., 1989. Surrogate endpoints in clinical trials: definition and operational criteria. Stat. Med. 8, 431–440. https://doi.org/10.1002/sim.4780080407.

Proix, T., Bartolomei, F., Guye, M., Jirsa, V.K., 2017. Individual brain structure and modelling predict seizure propagation. Brain 140, 641–654. https://doi.org/10.1093/brain/awx004.

Rehme, A.K., Volz, L.J., Feis, D.-L., Bomilcar-Focke, I., Liebig, T., Eickhoff, S.B., Fink, G. R., Grefkes, C., 2015. Identifying neuroimaging markers of motor disability in acute stroke by machine learning techniques. Cereb. Cortex N. Y. N 1991 (25), 3046–3056. https://doi.org/10.1093/cercor/bhu100.

Ren, C., Guingab-Cagmat, J., Kobeissy, F., Zoltewicz, S., Mondello, S., Gao, M., Hafeez, A., Li, N., Geng, X., Larner, S.F., Anagli, J., Hayes, R.L., Ji, X., Ding, Y., 2014. A neuroproteomic and systems biology analysis of rat brain post intracerebral hemorrhagic stroke. Brain Res. Bull. 102, 46–56. https://doi.org/10.1016/j.brainresbull.2014.02.005.

Richardson, W.S., Wilson, M.C., Nishikawa, J., Hayward, R.S., 1995. The well-built clinical question: a key to evidence-based decisions. ACP J. Club 123, A12–A13.

Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P., Van De Ville, D., 2011. Decoding brain states from fMRI connectivity graphs. NeuroImage, Multivariate Decoding and Brain Reading 56, 616–626. https://doi.org/10.1016/j.neuroimage.2010.05.081.

Ritter, P., Schirner, M., McIntosh, A.R., Jirsa, V.K., 2013. The Virtual Brain Integrates Computational Modeling and Multimodal Neuroimaging. Brain Connect. 3, 121–145. https://doi.org/10.1089/brain.2012.0120.

Röder, C.H., Hoogendam, J.M., van der Veen, F.M., 2010. FMRI, antipsychotics and schizophrenia. Influence of different antipsychotics on BOLD-signal. Curr. Pharm. Des. 16, 2012–2025. https://doi.org/10.2174/138161210791293088.

Sackett, D.L., Straus, S.E., Haynes, B.R., Rosenberg, W., Richardson, W.S., 2000. Evidence-Based Medicine: How to Practice and Teach EBM (Book with CD-ROM): Amazon.co.uk: David L. Sackett, Sharon E. Straus MD Dr., W. Scott Richardson MD Dr., William Rosenberg, R. Brian Haynes MD Dr.: 9780443062407: Books, 2. ed. Churchill Livingstone.

Salvalaggio, A., De Filippo De Grazia, M., Zorzi, M., Thiebaut de Schotten, M., Corbetta, M., 2020. Post-stroke deficit prediction from lesion and indirect structural and functional disconnection. Brain 143, 2173–2188. doi: 10.1093/brain/awaa156.

Schirner, M., Domide, L., Perdikis, D., Triebkorn, P., Stefanovski, L., Pai, R., Prodan, P., Valean, B., Palmer, J., Langford, C., Blickensdörfer, A., van der Vlag, M., Diaz-Pier, S., Peyser, A., Klijn, W., Pleiter, D., Nahm, A., Schmid, O., Woodman, M., Zehl, L., Fousek, J., Petkoski, S., Kusch, L., Hashemi, M., Marinazzo, D., Mangin, J.-F., Flöel, A., Akintoye, S., Stahl, B.C., Cepic, M., Johnson, E., Deco, G., McIntosh, A. R., Hilgetag, C.C., Morgan, M., Schuller, B., Upton, A., McMurtrie, C., Dickscheid, T., Bjaalie, J.G., Amunts, K., Mersmann, J., Jirsa, V., Ritter, P., 2022. Brain simulation as a cloud service: The Virtual Brain on EBRAINS. NeuroImage 251, 118973.

Shmuel, A., Leopold, D.A., 2008. Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: Implications for functional connectivity at rest. Hum. Brain Mapp. 29, 751–761. https://doi.org/10.1002/hbm.20580.

Silva, M.A., See, A.P., Essayed, W.I., Golby, A.J., Tie, Y., 2018. Challenges and techniques for presurgical brain mapping with functional MRI. NeuroImage Clin. 17, 794–803. https://doi.org/10.1016/j.nicl.2017.12.008.

Stephan, K.E., Mathys, C., 2014. Computational approaches to psychiatry. Curr. Opin. Neurobiol. 25, 85–92. https://doi.org/10.1016/j.conb.2013.12.007.

Tax, C.M., Grussu, F., Kaden, E., Ning, L., Rudrapatna, U., John Evans, C., St-Jean, S., Leemans, A., Koppers, S., Merhof, D., Ghosh, A., Tanno, R., Alexander, D.C., Zappalà, S., Charron, C., Kusmia, S., Linden, D.E., Jones, D.K., Veraart, J., 2019. Cross-scanner and cross-protocol diffusion MRI data harmonisation: a benchmark database and evaluation of algorithms. NeuroImage 195, 285–299. https://doi.org/10.1016/j.neuroimage.2019.01.077.

Tibben, A., Roos, R.a.C., Niermeijer, M.F., 1997. Psychological consequences of presymptomatic testing for Huntington's disease. The Lancet 349, 809. https://doi.org/10.1016/S0140-6736(05)60242-5.

Tscherpel, C., Dern, S., Hensel, L., Ziemann, U., Fink, G.R., Grefkes, C., 2020. Brain responsivity provides an individual readout for motor recovery after stroke. Brain 143, 1873–1888. https://doi.org/10.1093/brain/awaa127.

Varoquaux, G., Raamana, P.R., Engemann, D.A., Hoyos-Idrobo, A., Schwartz, Y., Thirion, B., 2017. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. NeuroImage, Individual Subject Prediction 145, 166–179. https://doi.org/10.1016/j.neuroimage.2016.10.038.

Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A.J., Alfaro-Almagro, F., Smith, S.M., Woolrich, M.W., 2018. Discovering dynamic brain networks from big data in rest and task. NeuroImage, Brain Connectivity Dynamics 180, 646–656. https://doi.org/10.1016/j.neuroimage.2017.06.077.

Wang, X.-J., Krystal, J.H., 2014. Computational psychiatry. Neuron 84, 638–654. https://doi.org/10.1016/j.neuron.2014.10.018.

Ward, J., Lyall, L.M., Bethlehem, R.A.I., Ferguson, A., Strawbridge, R.J., Lyall, D.M., Cullen, B., Graham, N., Johnston, K.J.A., Bailey, M.E.S., Murray, G.K., Smith, D.J., 2019. Novel genome-wide associations for anhedonia, genetic correlation with psychiatric disorders, and polygenic association with brain structure. Transl. Psychiatry 9, 327. https://doi.org/10.1038/s41398-019-0635-y.

Westeneng, H.-J., Debray, T.P.A., Visser, A.E., van Eijk, R.P.A., Rooney, J.P.K., Calvo, A., Martin, S., McDermott, C.J., Thompson, A.G., Pinto, S., Kobeleva, X., Rosenbohm, A., Stubendorff, B., Sommer, H., Middelkoop, B.M., Dekker, A.M., van Vugt, J.J.F.A., van Rheenen, W., Vajda, A., Heverin, M., Kazoka, M., Hollinger, H., Gromicho, M., Körner, S., Ringer, T.M., Rödiger, A., Gunkel, A., Shaw, C.E., Bredenoord, A.L., van Es, M.A., Corcia, P., Couratier, P., Weber, M., Grosskreutz, J., Ludolph, A.C., Petri, S., de Carvalho, M., Van Damme, P., Talbot, K., Turner, M.R., Shaw, P.J., Al-Chalabi, A., Chiò, A., Hardiman, O., Moons, K.G.M., Veldink, J.H., van den Berg, L.H., 2018. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. Lancet Neurol. 17, 423–433. https://doi.org/10.1016/S1474-4422(18)30089-9.

Wickenberg-Bolin, U., Göransson, H., Fryknäs, M., Gustafsson, M.G., Isaksson, A., 2006. Improved variance estimation of classification performance via reduction of bias caused by small sample size. BMC Bioinformatics 7, 127. https://doi.org/10.1186/1471-2105-7-127.

Wilson, R.C., Collins, A.G., 2019. Ten simple rules for the computational modeling of behavioral data. eLife 8, e49547.

Woo, C.-W., Chang, L.J., Lindquist, M.A., Wager, T.D., 2017. Building better biomarkers: brain models in translational neuroimaging. Nat. Neurosci. 20, 365–377. https://doi.org/10.1038/nn.4478.

Woolrich, M.W., Stephan, K.E., 2013. Biophysical network models and the human connectome. NeuroImage, Mapping the Connectome 80, 330–338. https://doi.org/10.1016/j.neuroimage.2013.03.059.

Worsley, K.J., Evans, A.C., Marrett, S., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. J. Cereb. Blood Flow Metab. Off. J. Int. Soc. Cereb. Blood Flow Metab. 12, 900–918. https://doi.org/10.1038/jcbfm.1992.127.

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4, 58–73. https://doi.org/10.1002/(SICI)1097-0193(1996)4:1<58::AID-HBM4>3.0.CO;2-O.

Yamashita, A., Yahata, N., Itahashi, T., Lisi, G., Yamada, T., Ichikawa, N., Takamura, M., Yoshihara, Y., Kunimatsu, A., Okada, N., Yamagata, H., Matsuo, K., Hashimoto, R., Okada, G., Sakai, Y., Morimoto, J., Narumoto, J., Shimada, Y., Kasai, K., Kato, N., Takahashi, H., Okamoto, Y., Tanaka, S.C., Kawato, M., Yamashita, O., Imamizu, H., 2019. Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. PLOS Biol. 17, e3000042.

Yao, D., Guo, X., Zhao, Q., Liu, L., Cao, Q., Wang, Y., D Calhoun, V., Sun, L., Sui, J., 2018. Discriminating ADHD From Healthy Controls Using a Novel Feature Selection Method Based on Relative Importance and Ensemble Learning. Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf. 2018, 4632–4635. doi: 10.1109/EMBC.2018.8513155.