



# NETWORK NEUROSCIENCE

an open access  journal



Citation: Karvelis, P., Charlton, C. E., Allohverdi, S. G., Bedford, P., Hauke, D. J., & Diaconescu, A. O. (2022). Computational approaches to treatment response prediction in major depression using brain activity and behavioral data: A systematic review. *Network Neuroscience*, 6(4), 1066–1103. [https://doi.org/10.1162/netn\\_a\\_00233](https://doi.org/10.1162/netn_a_00233)

DOI:  
[https://doi.org/10.1162/netn\\_a\\_00233](https://doi.org/10.1162/netn_a_00233)

Received: 14 October 2021  
Accepted: 14 January 2022

Competing Interests: The authors have declared that no competing interests exist.

Corresponding Author:  
Povilas Karvelis  
[povilas.karvelis@camh.ca](mailto:povilas.karvelis@camh.ca)

Handling Editor:  
Olaf Sporns

Copyright: © 2022  
Massachusetts Institute of Technology  
Published under a Creative Commons  
Attribution 4.0 International  
(CC BY 4.0) license



FOCUS FEATURE:  
Connectivity, Cognition, and Consciousness

## Computational approaches to treatment response prediction in major depression using brain activity and behavioral data: A systematic review

Povilas Karvelis<sup>1\*</sup> , Colleen E. Charlton<sup>1\*</sup>, Shona G. Allohverdi<sup>1</sup>, Peter Bedford<sup>1</sup>, Daniel J. Hauke<sup>1,2,3</sup> , and Andreea O. Diaconescu<sup>1,4,5,6</sup> 

<sup>1</sup>Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada

<sup>2</sup>Department of Psychiatry (UPK), University of Basel, Basel, Switzerland

<sup>3</sup>Department of Mathematics and Computer Science, University of Basel, Basel, Switzerland

<sup>4</sup>University of Toronto, Department of Psychiatry, Toronto, ON, Canada

<sup>5</sup>Institute of Medical Sciences, University of Toronto, Toronto, ON, Canada

<sup>6</sup>Department of Psychology, University of Toronto, Toronto, ON, Canada

\*These authors contributed equally to this work.

**Keywords:** Major depressive disorder, Treatment response prediction, Machine learning, Computational psychiatry, fMRI, EEG

### ABSTRACT

Major depressive disorder is a heterogeneous diagnostic category with multiple available treatments. With the goal of optimizing treatment selection, researchers are developing computational models that attempt to predict treatment response based on various pretreatment measures. In this paper, we review studies that use brain activity data to predict treatment response. Our aim is to highlight and clarify important methodological differences between various studies that relate to the incorporation of domain knowledge, specifically within two approaches delineated as data-driven and theory-driven. We argue that theory-driven generative modeling, which explicitly models information processing in the brain and thus can capture disease mechanisms, is a promising emerging approach that is only beginning to be utilized in treatment response prediction. The predictors extracted via such models could improve interpretability, which is critical for clinical decision-making. We also identify several methodological limitations across the reviewed studies and provide suggestions for addressing them. Namely, we consider problems with dichotomizing treatment outcomes, the importance of investigating more than one treatment in a given study for differential treatment response predictions, the need for a patient-centered approach for defining treatment outcomes, and finally, the use of internal and external validation methods for improving model generalizability.

### AUTHOR SUMMARY

Individuals with major depressive disorder (MDD) vary in their response to available treatments, rendering treatment selection a challenging task. In this paper, we review studies applying computational models for predicting treatment response in MDD based on measures of brain activity. We discuss methodological differences across studies, focusing on how they incorporate existing knowledge about MDD and how that affects interpretability of model predictions. In this context, we argue that theory-driven generative modeling,

which explicitly models information processing in the brain and thus can capture disease mechanisms, is a promising emerging approach for treatment response prediction. Finally, we identify several other important limitations that are holding back the translation of these tools into clinical practice.

---

## INTRODUCTION

Depressive disorders are the third highest cause of years lived with disability, affecting more than 264 million people worldwide (James et al., 2018). Major depressive disorder (MDD) is the most frequent type of depressive disorder (Vandeleur et al., 2017) and is characterized by depressed mood, diminished interests or pleasure, vegetative symptoms (e.g., appetite or sleep disturbances), and impaired cognition (e.g., feelings of worthlessness or inappropriate guilt; American Psychiatric Association, 2013). A key challenge in the treatment of MDD is the heterogeneity of illness course and treatment response (Luedtke & Kessler, 2021). Patients often show diverse initial symptoms with divergent disease trajectories over time, and some symptoms persist in spite of treatment (Rush et al., 2006). As a result, many patients face a long and painful trial-and-error process to identify the right treatment.

A promising way forward is to leverage computational models for understanding the heterogeneity of MDD and identifying individual predictors of differential treatment response. Broadly defined, computational psychiatry aims to formalize the relationship between the brain's neurobiology, its environment, and psychiatric symptoms in computational terms (Adams, Huys, & Roiser, 2016). Within computational psychiatry, there are two conceptually different approaches: data-driven and theory-driven modeling (Huys, Maia, & Frank, 2016; Stephan et al., 2017). **Data-driven approaches** are domain-knowledge-agnostic, and classical statistics or machine learning techniques are used for exploratory analyses to discover predictive patterns in high-dimensional data. Conversely, **theory-driven approaches** use models that rely on domain-knowledge-derived hypotheses about the processes underlying neural and/or behavioral data. While advances in each approach have helped progress clinical research, each method comes with different trade-offs. Moreover, these approaches are not mutually exclusive and are often integrated in different ways. With this in mind, we review electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) studies on MDD treatment response prediction and organize methodological differences among studies along the data-driven versus theory-driven dimensions. We limit our review to studies that predict treatment response using out-of-sample testing, while studies that report simple (in-sample) associations between the variables and treatment response outcomes will not be discussed.

In what follows, we first provide a brief introduction into data-driven and theory-driven approaches. We propose that the data-driven versus theory-driven trade-off could be seen along two dimensions related to the way data are collected and the way data are processed. This provides background for the following section, where we briefly summarize fMRI and EEG studies using data- or theory-driven methods for treatment response prediction in MDD, going from the least theory-driven to the most theory-driven approaches. Next, we present several promising theory-driven developments that are only beginning to be utilized for treatment response prediction in MDD. Finally, we discuss more generally what challenges are remaining for these tools to be translated into clinical practice.

## DIFFERENT TYPES OF APPROACHES

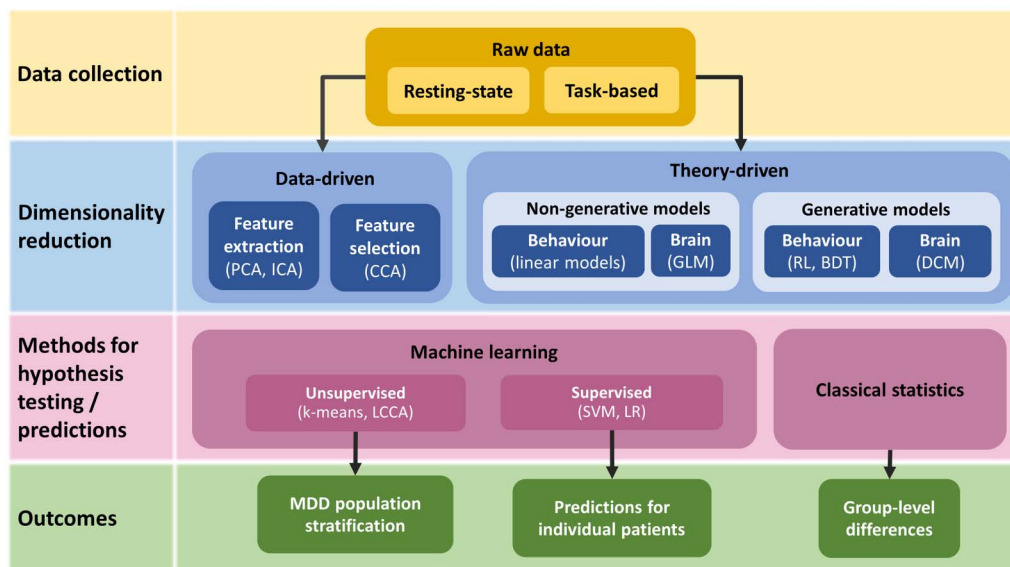
### Data-Driven Approaches

Data-driven approaches aim to identify patterns or predictive relationships in high-dimensional datasets without relying directly on domain knowledge (Figure 1)—such as predicting treatment response based on neuroimaging data without requiring specific hypotheses about the neural mechanisms underlying MDD. To obtain discriminative features from the raw data, such approaches typically use either feature extraction or feature selection methods. Feature extraction involves constructing a smaller set of new features from the existing ones by applying techniques such as independent component analysis (ICA) or principal component analysis (PCA). Feature selection, on the other hand, does not transform the original features, but simply removes those that are irrelevant, such as highly intercorrelated features or features that do not correlate with the target variable (e.g., treatment response).

Supervised learning:

A set of techniques for inferring a function that maps input features (e.g., neuroimaging features) to output labels (e.g., treatment response) based on training examples of feature-label pairs.

The resulting features are then used to train machine learning algorithms to predict treatment response. In this review, we consider two main types of machine learning algorithms: supervised and unsupervised. In **supervised learning**, an algorithm learns a function that maps input features (e.g., neuroimaging measures) to output labels (e.g., treatment outcome). Supervised algorithms are used for classification, the prediction of a discrete label (e.g., remission vs.



**Figure 1.** A conceptual overview illustrating common methods employed by data-driven and theory-driven approaches to study treatment response prediction in MDD. At the data collection stage, theory-driven knowledge can be incorporated via task-based designs, which probe specific cognitive functions. Most often, resting-state data will be processed with data-driven methods by performing feature selection, for example, using canonical correlation analysis (CCA), or feature extraction, for example, using principal or independent component analysis (PCA or ICA). Theory-driven dimensionality reduction is most often applied to task-based data using non-generative models, such as linear models used for obtaining summary statistics from behavioral data or the commonly used general linear model (GLM) in neuroimaging data analysis. A more advanced approach is to use generative models of behavior, such as reinforcement learning (RL) or Bayesian decision theory (BDT) models, which can be fit to behavioral data, or dynamic causal modeling (DCM), which can be applied to neuroimaging data. Next, the obtained features are used to train machine learning algorithms. Using supervised learning, such as support vector machines (SVM) or logistic regression (LR), one can determine the predictive ability of these features at the level of an individual patient. Unsupervised learning, such as k-means or latent class cluster analysis (LCCA), on the other hand, is primarily focused on stratifying MDD patient population, but the determined subpopulations could subsequently inform treatment response prediction. In contrast to machine learning approaches, classical statistics methods that are concerned with uncovering group-level effects (e.g., group differences between responders and nonresponders) and that do not provide predictions at the level of an individual patient are thus not included in this review.

Unsupervised learning:  
A set of techniques for finding patterns in the distribution of features (e.g., neuroimaging measures) without using labels (e.g., treatment response).

nonremission), or regression, the prediction of a continuous variable (e.g., symptom improvement after treatment). Popular supervised classifiers include support vector machines (SVM; Cortes & Vapnik, 1995), logistic regression (Wright, 1995), and decision trees (e.g., Breiman, 2001). Comparatively, **unsupervised learning** does not require output labels, but instead finds patterns in the distribution of the input features. Clustering, the process of grouping data together based on underlying similarities, may be used to identify MDD subtypes or treatment response profiles, although ascertaining the clinical usefulness of the discovered clusters may require additional analyses with the data labels. As we will see in section *Studies Examining Treatment Response Prediction in MDD*, most of the neuroimaging studies on treatment response prediction in MDD to date have employed supervised learning methods.

### Theory-Driven Approaches

Theory-driven approaches employ domain-knowledge-informed techniques that utilize hypotheses about the underlying mechanisms of MDD (Figure 1). The models used in this approach can broadly be divided into generative and non-generative models. Non-generative models would subsume linear models used for obtaining summary statistics from behavioral data or the commonly used general linear model (GLM) in neuroimaging data analysis. Generally, such methods are used for exploratory analyses or to test hypotheses in terms of a specific contrast-, condition-, or group-related effect. In contrast, **generative models** aim to explicitly describe a mechanism that underlies neural or behavioral data in computational terms and thus are able to generate such data. These methods can span multiple levels of analyses, from biophysically informed models that describe the dynamics at a single neuron (e.g., ion channel conductances, membrane potential, firing rate), to generative models of brain responses that investigate experimentally induced coupling among brain areas, to generative models of behavior that describe information processing underlying decision-making. This type of approach allows for inference on disease mechanisms and provides a detailed model of the disease (Huys et al., 2016; Stephan et al., 2017). Another notable advantage of generative models is that they can be used to investigate the behavior of the system under different conditions through simulations. In this manner, simulations can be employed to generate new hypotheses about disease mechanisms and achieve a better understanding of the neurobiology of treatment effects (Stephan et al., 2017).

Through various (mostly non-generative) theory-driven approaches, MDD has been associated with deficits in reward processing and emotion regulation (Han, Ham, & Kim, 2021; Phillips et al., 2015), implicating serotonergic and dopaminergic neurocircuits, respectively (Keren et al., 2018; Kupfer, Frank, & Phillips, 2012; Pizzagalli, 2014). Some of the main findings involve elevated amygdala activation in response to negative emotional stimuli, increased activity of the anterior cingulate/ventromedial prefrontal cortex (ACC/vmPFC), which are involved in automatic emotion regulation, and attenuated activity of the dorsolateral PFC (dlPFC), which is involved in voluntary emotion regulation (Han et al., 2021; Phillips et al., 2015). With regards to reward processing, decreased activity of the ventral striatum has emerged as the most prominent finding (Keren et al., 2018). Large-scale network studies have also revealed increased connectivity within the affective network, reduced connectivity within the frontal-striatal reward network, diminished connectivity within the central executive network, and hyperconnectivity of the default mode network (DMN; Kaiser, Andrews-Hanna, Wager, & Pizzagalli, 2015; Li et al., 2018). Importantly, changes in each of these networks have been related to different symptoms (e.g., anhedonia being more associated with the reward network, and rumination with the DMN) and thus might differentially respond to treatments (Chahal, Gotlib, & Guyer, 2020; Li et al., 2018). Despite the increasing understanding of

the neural mechanisms underlying MDD, recent review articles on regional activation and functional connectivity measures suggest, however, a lack of reliable neuromarkers for treatment response prediction (Fonseka, MacQueen, & Kennedy, 2018; Taylor, Kurt, & Anand, 2021), which highlights the need for methodological advancements.

### Combining Data-Driven and Theory-Driven Approaches

Data-driven and theory-driven approaches can be combined in different ways. This can be done along two dimensions: data collection and dimensionality reduction (Figure 1). In the former case, theory-driven knowledge is incorporated into the decision about what type of data is collected, for example, by probing cognitive mechanisms known to be implicated in MDD with cognitive tasks. In the latter case, theory-driven knowledge is incorporated by selecting a subset of data features (e.g., brain regions) or deriving the features based on previous literature or a priori hypotheses.

One of the main benefits of incorporating more theory-driven elements into treatment response prediction is the **interpretability** of the discriminative features (Finn, 2021; Stephan et al., 2017). Increased interpretability can help both clinicians and researchers better understand how model predictions and the associated neural markers relate to the behavioral and cognitive symptoms experienced by patients. For clinicians, interpretability is essential for making informed judgements based on a model's prediction (Kelly, Karthikesalingam, Suleyman, Corrado, & King, 2019; Rudin, 2019). For researchers, interpretability can help them understand why certain treatments do not work for some individuals—which can provide insights for the development of new treatments (Stephan et al., 2017).

At the stage of data collection, interpretability can be improved by using **task-based neuroimaging** paradigms. Compared with **resting-state neuroimaging** protocols (which are more common in data-driven approaches), task-based paradigms are usually more sensitive to brain–behavior and brain–mind relationships (Finn, 2021). While resting-state data collection could also be motivated by specific hypotheses about rumination and mind wandering in MDD, and thus have a theory-driven motivation, without sufficient experimental control (e.g., experience sampling or retrospection) to relate the recorded brain activity to experience, treatment response predictions derived from such data can be difficult to interpret.

Interpretability can be further improved at the dimensionality reduction stage by applying theory-driven techniques, which incorporate knowledge or hypotheses about mechanisms underlying MDD. This is in contrast to data-driven dimensionality reduction, which does not require such knowledge and, as a result, offers less interpretable features. State-of-the-art theory-driven dimensionality reduction involves the use of generative models that describe information processing dynamics in the brain and by doing so is able to capture disease mechanisms in greater detail (Figure 1). Applying machine learning classification or clustering methods using features derived from fitting such generative models to data has been referred to as **generative embedding** (Brodersen et al., 2011; Frässle et al., 2018). However, this technique is only beginning to be utilized for treatment response prediction in MDD and is yet to prove its potential (see the section *Generative Modeling and Generative Embedding*).

### STUDIES EXAMINING TREATMENT RESPONSE PREDICTION IN MDD

In this section, we briefly summarize recent studies on treatment response prediction in MDD, going from the least to the most theory-driven approaches (Tables 1 and 2). Our aim is to highlight the variety of methodologies used and to draw attention to the important distinctions of data- versus theory-driven strategies (Figure 1). We restrict our review to studies that performed

#### Interpretability:

The degree to which a user can comprehend why certain predictions have been made by a predictive model.

#### Task-based neuroimaging:

Measures brain activity that is functionally involved in a specific cognitive or behavioral task.

#### Resting-state neuroimaging:

Measures brain activity and regional interactions that occur in the absence of a stimulus or a task (“at rest”).

#### Generative embedding:

The use of features extracted via a generative model as input for supervised or unsupervised machine learning algorithms.

**Table 1.** Overview of data-driven MDD treatment response prediction studies. Relevant model performance metrics (BAC and  $R^2$ ) that were not reported by the studies but were possible to calculate from the reported values are included in parentheses. MDD - major depressive disorder, TRD - treatment-resistant depression, NDRI - norepinephrine-dopamine reuptake inhibitor, SSRI - selective serotonin reuptake inhibitor, SNRI - serotonin-norepinephrine reuptake inhibitor, rTMS - repetitive transcranial magnetic stimulation, tDCS - transcranial direct-current stimulation, ECT - electroconvulsive therapy, MADRS - Montgomery-Asberg Depression Rating Scale, HAMD - Hamilton Rating Scale for Depression, BDI - Beck Depression Inventory, QIDS-SR - Quick Inventory of Depressive Symptomatology—Self Report, ROI - region of interest, FC - functional connectivity, ICA - independent component analysis, SVM - support vector machine, LDA - linear discriminant analysis, CV - cross-validation, LOOCV - leave-one-out cross-validation, LOSOCV - leave-one-site-out cross-validation, Acc - accuracy, SE - sensitivity, SP - specificity, BAC - balanced accuracy, RMSE - root mean square error, dmPCF - dorsomedial prefrontal cortex, ACC - anterior cingulate cortex.

Reference	Subjects	Treatment	Modality	Outcome definition	Features	Model	Validation	Performance
Jaworska et al. (2019)	51 MDD	NDRI (bupropion), SSRI (escitalopram), or combination of both	rsEEG	Response: $\geq 50\%$ ↓ in MADRS	Demographics, baseline, & Week 1 clinical data, EEG power features, current source density	Random forest	10-fold CV	Acc 88% SE 77% SP 99%
Zhdanov et al. (2020)	122 MDD	SSRI (escitalopram)	rsEEG	Response: $\geq 50\%$ ↓ in MADRS	Electrode-level & source-level spectral features, multiscale entropy-based & microstate-based features	SVM	Leave-one-site-out CV (LOSOCV)	BAC 79% SE 67% SP 91%
Khodayari-Rostamabad et al. (2010)	22 MDD	SSRI (mainly sertraline)	rsEEG	Response: $\geq 25\%$ ↓ in HAMD-17	Spectral coherence, mutual information between electrode pairs, absolute & relative power spectral density	Kernel partial least squares regression	Nested CV	Acc 87% SE 88% SP 86%
Khodayari-Rostamabad et al. (2013)	22 TRD	SSRI (sertraline, citalopram, fluvoxamine, or paroxetine)	rsEEG	Response: $\geq 30\%$ ↓ in HAMD-17	Power spectral density, squared spectral coherence, mutual information, left-to-right hemispheres, & anterior/posterior power ratio	Mixture of factor analysis	k-fold CV	Acc 88% SE 95% SP 81%
Rabinoff et al. (2011)	25 MDD	SSRI (fluoxetine) or SNRI (venlafaxine)	rsEEG	Response: HAMD-17 $\leq 10$	Absolute & relative power, cordance features	Classification and regression trees (CART)	LOOCV	BAC 93% SE 85% SP 100%

Table 1. (continued)

Reference	Subjects	Treatment	Modality	Outcome definition	Features	Model	Validation	Performance
Shahabi, Shalhaf, and Maghsoudi (2021)	30 MDD	SSRI (type not specified)	rsEEG	Response: $\geq 50\%$ $\downarrow$ in BDI	3D images constructed from EEG signal	Convolutional neural networks	10-fold CV	Acc 97% SE 96% SP 97%
W. Wu et al. (2020)	109 MDD (sertraline), 119 MDD (placebo)	SSRI (sertraline)	rsEEG	$\Delta$ in HAMD-17	Theta, alpha, beta, gamma band power of latent signal	Linear regression	10-fold CV	$(R^2 = 0.36)$ $r = 0.60$ RMSE = 5.68 $p = 2.88 \times 10^{-11}$
Rajpurkar et al. (2020)	518 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	rsEEG	$\Delta$ in HAMD-21 (individual symptoms)	Absolute & relative power of delta, theta, alpha, beta, & gamma frequency bands in frontal & occipital regions	Gradient-boosted decision trees (GBDT)	5-fold stratified CV	Concordance index of $\geq 0.8$ on 12 out of 21 symptoms $R^2$ 0.3–0.7
Khodayari-Rostamabad, Reilly, Hasey, de Bruin, and MacCrimmon (2011)	27 TRD	rTMS	rsEEG	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Anterior/posterior power ratios at various frequencies	Mixture of factor analysis	k-fold CV	Acc 80% (BAC 81%) SE 78% SP 83%
N. Bailey et al. (2019)	50 TRD	rTMS	rsEEG	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Mood features, theta & alpha power & connectivity, frontal theta cordance & alpha peak frequency	SVM	5-fold CV	BAC 86% SE 84% SP 89%
Hasanzadeh et al. (2019)	46 MDD	rTMS	rsEEG	Response: $\geq 50\%$ $\downarrow$ in HAMD-17 or BDI-II	Nonlinear, power spectral density, bispectrum, frontal & prefrontal cordance	k-nearest neighbors	LOOCV	BAC 91% SE 87% SP 96%
Al-Kaysi et al. (2017)	10 MDD	tDCS	rsEEG	Response: $\geq 50\%$ $\downarrow$ in MADRS	Power spectral density in delta, theta, alpha, beta, & gamma frequency bands	SVM, LDA, extreme learning machine	LOOCV	Mood Labels Channels FC4-AF8: Acc 76%; Cognition Labels Channels CPz-CP2: Acc 92%

Tian et al. (2020)	106 MDD	SSRI (escitalopram)	rsfMRI	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Multilayer modularity framework applied to the whole brain to obtain measures of functional integration & segregation among 95 ROIs	SVM	Leave-one-site-out CV (LOSOCV)	BAC 71%
Klöbl et al. (2020)	29 MDD	SSRI (escitalopram)	rsfMRI	$\Delta$ in HAMD-17; Response: $\geq 50\%$ $\downarrow$ in HAMD-17; Remission: HAMD-17 $\leq 7$	Whole-brain FC	Linear regression	k-fold CV	HAMD-sum: $r = 0.51$ Response: BAC 60%, AUC 68% Remission: BAC 68%, AUC 73%
Chin Fatt et al. (2020)	132 MDD (sertraline), 132 MDD (placebo)	SSRI (sertraline)	rsfMRI	$\Delta$ in HAMD-17	Cortical & subcortical seed-based FC	Linear mixed model	LOOCV	( $R^2 = 0.05-0.13$ ) $r = 0.22-0.36$
Korgaonkar et al. (2020)	163 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	rsfMRI	Remission: HAMD-17 $\leq 7$	Whole-brain network intrinsic FC	Logistic regression	Hold-out test set	Average connectivity measures: Acc 69% (BAC 67%) SE 58% SP 76% Individual network connectivity: Acc 69% (BAC 68%) SE 63% SP 72%



Table 1. (continued)

Reference	Subjects	Treatment	Modality	Outcome definition	Features	Model	Validation	Performance
Nemati et al. (2020)	99 MDD (sertraline), 103 MDD (placebo), & 19 MDD (ketamine), 19 MDD (active control), 18 MDD (inactive control)	SSRI (sertraline), ketamine	rsfMRI	$\Delta$ in HAMD-17	Network restricted connectivity	Network restricted strength predictive (linear) model	10-fold CV	Sertraline (vs. placebo): $r = 0.27$ ( $R^2 = 0.07$ ), $p = 0.003$ ; Ketamine (vs. active placebo): $r = 0.57$ ( $R^2 = 0.32$ ), $p = 0.0002$
Fan et al. (2020)	97 MDD (sertraline), 103 MDD (placebo)	SSRI (sertraline)	rsfMRI	$\Delta$ in HAMD-17 (%)	Network restricted connectivity	Network restricted strength predictive (linear) model	10-fold CV	Response to sertraline or placebo: ( $R^2 = 0.04$ ) $r = 0.19$ , $p = 0.03$
Ju et al. (2020)	108 MDD	Various drugs; primarily: paroxetine, other SSRIs, sedative hypnotics, NDRI (bupropion)	rsEEG	$\Delta$ in HAMD-24	Whole-brain FC matrices	Connectome-based predictive modeling	LOOCV	$r = 0.43$ ( $R^2 = 0.19$ ), $p = 2.73 \times 10^{-6}$
Kong et al. (2021)	82 MDD	Antidepressants (type not specified)	rsfMRI	Response: $>50\%$ $\downarrow$ in HAMD-21	Dynamic functional networks	Spatiotemporal graph convolutional network	10-fold CV	Acc 90% (BAC 89%) SE 85% SP 93%
Drysdale et al. (2017)	154 MDD	rTMS	rsfMRI	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Whole-brain FC matrices & biotype diagnosis	SVM	LOOCV	Only FC feature: Acc 78.3% FC features & biotype diagnosis: Acc 89.6%
Van Waarde et al. (2015)	45 severe/TRD	ECT	rsfMRI	Remission: MADRS score $\leq 10$	Standard group ICA extracted 25 rs-networks. Each network was used to train a classifier	SVM	LOOCV	Two rs-networks had significant accuracy: <b>dmPFC</b> : BAC 85%; SE 84%; SP 85%; <b>ACC</b> : BAC 78%; SE 80%; SP 75%

Leaver et al. (2018)	46 TRD	ECT	rsfMRI, sMRI, arterial spin labeled fMRI	Response: average % improvement in HAMD-17, MADRS, and QIDS-SR. Split point was 42.2% reduction.	Mean voxelwise cerebral blood flow, regional homogeneity, fractional amplitude of low-frequency fluctuations, gray matter volume	SVM	Nested CV	BAC 58–68% SE 54–64% SP 55–74%
Sun et al. (2020)	122 MDD or bipolar disorder	ECT	rsfMRI	$\Delta$ in HAMD-17; Remission: HAMD-17 score < 7	Negatively & positively correlated FC networks based on whole-brain rsFC	Linear regression	LOOCV	Negative FC networks: $r = 0.51$ ( $R^2 = 0.26$ ) Acc 76% (BAC 72%) SE 51% SP 92%

**Table 2.** Overview of data- and theory-driven MDD treatment response prediction studies. Relevant model performance metrics (BAC and  $R^2$ ) that were not reported by the studies but were possible to calculate from the reported values are included in parentheses. MDD - major depressive disorder, TRD - treatment-resistant depression, LLD - late-life depression, SSRI - selective serotonin reuptake inhibitor, rTMS - repetitive transcranial magnetic stimulation, SNRI - serotonin-norepinephrine reuptake inhibitor, NDRI - norepinephrine-dopamine reuptake inhibitor, CBT - cognitive behavioral therapy, ECT - electroconvulsive therapy, BDI - Beck Depression Inventory, HAMD - Hamilton Rating Scale for Depression, MADRS - Montgomery-Asberg Depression Rating Scale, QIDS-SR - Quick Inventory of Depressive Symptomatology—Self Report, SOFAS - Social and Occupational Functioning Assessment Scale, CIDI - Composite International Diagnostic Interview, LCI - Life Chart Interview, STFT - short-time Fourier transform, EMD - empirical mode decompositions, ACC - anterior cingulate cortex, FC - functional connectivity, ROI - region of interest, ICA - independent component analysis, BOLD - blood oxygen level dependent, dlPFC - dorsolateral prefrontal cortex, amPFC - anterior medial prefrontal cortex, mPFC - medial prefrontal cortex, DMN - default mode network, SN - salience network, PCC - posterior cingulate cortex, AN - affective network, VIS - visual, SVM - support vector machine, ROC - receiver operating characteristic, CV - cross-validation, LOOCV - leave-one-out cross-validation, Acc - accuracy, SE - sensitivity, SP - specificity, BAC - balanced accuracy, AUC - area under curve, RMSE - root mean square error.

Reference	Subjects	Treatment	Modality	Outcome definition	Features	Model	Validation	Performance
Mumtaz et al. (2017)	34 MDD	SSRIs (type not specified)	tbEEG: 3-stimulus visual oddball task	Response: $\geq 50\%$ $\downarrow$ in BDI-II	Combination of wavelet coefficients, STFT, & EMD features	Logistic regression	10-fold CV	Acc 92% SE 90% SP 90%
N. Bailey et al. (2018)	39 TRD	rTMS	tbEEG; Working memory task	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Baseline and Week 1 MADRS scores, task accuracy & reaction time, alpha, theta, gamma power & connectivity, theta gamma coupling	SVM	5-fold CV	BAC 91% SE 90% SP 92%
Miller et al. (2013)	17 MDD	SSRI (escitalopram)	tbfMRI; emotional words task	$\Delta$ in HAMD-24	Clusters whose active during negative word processing was associated with treatment outcome	Linear regression	10-fold CV	( $R^2 = 0.23$ ) $r = 0.48$ , $p < 0.05$
Godlewska et al. (2018)	32 MDD	SSRI (escitalopram)	tbfMRI; emotional faces task	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Mean cluster activity within ACC for sad vs. happy faces	Single-feature ROC	LOOCV	BAC 72%
Fu et al. (2008)	19 MDD	SSRI (fluoxetine)	tbfMRI; implicit sad facial affect recognition task	Remission: HAMD-17 $\leq 8$	Whole-brain FC for each facial expression intensity (low, medium, high)	SVM	LOOCV	Low intensity of sad facial expressions: SE 75% SP 62% $p$ value = 0.11 (n.s.)
Fonzo et al. (2019)	115 MDD (sertraline), 122 MDD (placebo)	SSRI (sertraline)	tbfMRI; Emotional conflict task	$\Delta$ in HAMD-17	incongruent trials—congruent trials brain activation in several ROIs	Relevance vector machine	10 $\times$ 10 stratified CV	( $R^2 = 0.24$ ) $r = 0.49$ , $p < 0.001$

Karim et al. (2018)	49 LLD	SNRI (venlafaxine)	rsfMRI and tbfMRI; emotional regulation & emotional reactivity task	Remission: MADRS $\leq 10$	Active regions during emotional reactivity, emotion regulation, or centrality at baseline & after single dose	Logistic regression	10-fold CV	(BAC 70%) AUC 77% SE 72% SP 68%
Williams et al. (2015)	80 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	tbfMRI; Supraliminal & subliminal facial emotion task	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Emotion vs. neutral amygdala activation	Discriminant analysis	LOOCV	All medication types: Acc 75%  SNRI only: Acc 77%
Goldstein-Piekarski et al. (2016)	70 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	tbfMRI; emotional faces task	Functional remission: HAMD-17 $\leq 7$ and QIDS-SR $\leq 5$ and $\geq 10$ improvement to achieve $\geq 61$ on SOFAS	Early life stress & amygdala reactivity	Logistic regression	LOOCV	(BAC 77%) SE 84% SP 69% AUC 81%
Crane et al. (2017)	29 MDD	SSRI (escitalopram) or SNRI (duloxetine)	tbfMRI; Go/No-Go task	Remission: HAMD score $< 8$	No-Go accuracy, two ICA component beta weights, & within-component clusters	Logistic regression, Random forest	LOOCV	Logistic regression: Acc 90% (BAC 90%) SE 90% SP 89% Random forest: Acc 84% BAC (82%) SE 84% SP 80%

Table 2. (continued)

Reference	Subjects	Treatment	Modality	Outcome definition	Features	Model	Validation	Performance
Tozzi et al. (2020)	124 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	tbfMRI; Go/No-Go task	Response: >50% ↓ in QIDS-SR-16	ROIs with BOLD response in No-Go > Go condition in MDD patients & healthy controls	Logistic regression	LOOCV	Venlafaxine response: BAC 79% SE 67% SP 89%; Sertraline response: BAC 84% SE 95% SP 74%
Marquand, Mourão-Miranda, Brammer, Cleare and Fu (2008)	20 MDD	SSRI (fluoxetine)	tbfMRI; n-back task	Response: ≥50% ↓ in HAMD-17	Principal components based on whole-brain activity for each task condition	SVM	LOOCV	BAC 69% SE 85% SP 52%
Meyer et al. (2019)	22 MDD	SSRI (escitalopram)	tbfMRI; n-back task	Remission: MADRS ≤ 5 Nonremission: MADRS ≥ 10	dIPFC, amPFC, & parietal lobe	Single-feature ROC	LOOCV	(BAC 87%) AUC 85% SE 82% SP 91%
Nguyen et al. (2019)	37 MDD	NDRI (bupropion)	tbfMRI; reward processing task	Δ in HAMD-17; Remission: HAMD-17 score < 7	ROIs from anticipation contrast maps & reward expectation contrast maps	Dense feedforward neural networks	3 × 3 nested CV	R <sup>2</sup> = 0.26 RMSE = 4.71 AUC 71%
Brandt et al. (2021)	90 MDD	SSRI (escitalopram); an optional switch to SNRI (duloxetine) from Week 4	tbfMRI; reward processing task	HAMD-6 Remission: >50% ↓ at week 4 <i>and</i> <5 at Week 8; Nonresponse: <25% ↓ a Week 4 <i>and</i> <50% ↓ at Week 8	Age, sex, baseline HAMD-6 and number of omissions in the task, & reward-related brain responses in striatum, anterior insula, & mPFC	Logistic regression	5-fold CV	AUC 56%
Costafreda, Khanna, Mourao-Miranda, and Fu (2009)	16 MDD	CBT	tbfMRI; sad facial expression task	Remission: HAMD-17 ≤ 7	Principal components based on whole-brain activity to viewing sad faces	SVM	LOOCV	BAC 79% SE 71% SP 86%

Cook et al. (2020)	129 MDD	SSRI (escitalopram) or NDRI (bupropion)	rsEEG	Remission: HAMD-17 $\leq 7$	Theta and alpha power values from FT7-FPz and FT8-FPz channels at baseline and Week 1 (QEEG)	Single-feature ROC analysis	Jack-knife CV	Escitalopram remission: Acc 64% (BAC 65%) SE 74% SP 55%
Erguzel et al. (2015)	55 TRD	rTMS	rsEEG	Response: $\geq 50\%$ $\downarrow$ in HAMD-17	Frontal cordance calculated from delta & theta bands (QEEG)	Artificial neural network	k-fold CV	Acc 85–89% (BAC 86–89%) SE 87–94% SP 84% AUC 87–91%
Patel et al. (2015)	33 LLD	SSRI (escitalopram) or SNRIs (duloxetine or venlafaxine)	rsfMRI; sMRI	Response: HAMD score $< 10$	Demographics, cognitive ability scores, functional connectivity index of dorsal DMN and anterior SN, & structural imaging variables	Alternating decision tree (ADTree)	Nested LOOCV	Acc 89% SE 89% SP 90%
Pei et al. (2020)	98 MDD	SSRIs (mainly escitalopram) & SNRIs (mainly venlafaxine)	rsfMRI	Response: $\geq 50\%$ $\downarrow$ in HAMD-6	14 priori brain regions of interest based on previous literature	SVM	LOOCV	BAC 81% SE 78% SP 84%
Sikora et al. (2016)	29 MDD	1-week placebo & 10-week open-label antidepressant (SSRI, SNRI, NDRI, atypical)	rsfMRI	$\Delta$ in QIDS	Baseline SN rsFC	Multivariate relevance vector regression	LOOCV	Placebo response: ( $R^2 = 0.17$ ) $R = 0.41$ , $p$ value = 0.018; Antidepressant response: $r = 0.03$ , $p$ value = 0.340

Table 2. (continued)

Reference	Subjects	Treatment	Modality	Outcome definition	Features	Model	Validation	Performance
Braund et al. (2022)	229 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	rsfMRI	Response: $\geq 50\%$ $\downarrow$ in HAMD-17 or QIDS-SR	Whole-brain network intrinsic FC associated with neuroticism	SVM	LOOCV	Acc 75% (BAC 74%) SE 63% SP 85% AUC 76%
Goldstein-Piekarski et al. (2018)	75 MDD	SSRI (escitalopram, sertraline) or SNRI (venlafaxine)	rsfMRI	Remission: HAMD-17 $\leq 7$	Seed-based posterior cingulate cortex	Logistic regression	LOOCV	PCC-ACC/mPFC: (BAC 78%) SE 73% SP 82% AUC 77%
C. T. Wu et al. (2018)	22 TRD	Sham-rTMS	rsfMRI	$\Delta$ in HAMD-17	Two feature sets: 1. Global brain activity; 2. Rostral ACC seed-based	Elastic-net regression	LOOCV	Global brain activity: ( $R^2 = 0.24$ ) $r = 0.49$ ; $p = 0.023$ ; rostral ACC FC: ( $R^2 = 0.25$ ) $r = 0.50$ ; $p = 0.018$
Cash et al. (2019)	47 MDD	rTMS	rsfMRI	Response: $\geq 25\%$ $\downarrow$ in MADRS; $\Delta$ in MADRS	rsFC within the DMN & AN	SVM	LOOCV, k-fold CV	Acc 93% SE 95% SP 92 % ( $R^2 = 0.46$ ) $r = 0.68$ $p$ value < 0.001
Hopman et al. (2021)	70 TRD	rTMS	rsfMRI	Response: $\geq 50\%$ $\downarrow$ in MADRS	Seed-based analysis with left dlPFC and subgenual ACC	SVM	Nested CV	Acc 89%
Moreno-Ortega et al. (2019)	18 TRD	ECT	rsfMRI	Remission: 24-HAMD score $\leq 7$	dlPFC, DMN, & VIS networks	Logistic regression	LOOCV	Acc 89%

Frässle et al. (2020)	85 MDD	SSRIs or no treatment	tbfMRI; emotional face perception task	Long-term outcomes: remitted vs. improved vs. chronic, based on a latent class growth analysis using CIDI and LCI	Effective connectivity pattern in the network mediating emotional face perception	SVM	Nested LOOCV	Chronic vs. remitted:  AUC 87%  BAC 79%  SE 97%  SP 60%  Improved vs. remitted:  AUC 63%  BAC 61%  SE 77%  SP 45%
Queirazza et al. (2019)	25 MDD	Computerized CBT	model-based fMRI; RL task	Response: >50% ↓ in BDI-II	Neural activity encoding weighted reward prediction errors	SVM	LOOCV	AUC 82%  BAC 72%  SE 63%  SP 83%



Cross-validation:

A validation technique used to estimate model performance by training the model on a subsample of data and validating model predictions on the remaining sample. This process is often repeated by permuting the data.

External validation:

A technique that validates a model trained on one dataset on an independently acquired dataset to determine the model's generalizability and reproducibility.

explicit analysis of treatment response prediction and incorporated validation techniques (e.g., **cross validation** or **external validation**) to increase the generalizability of findings. Studies that investigated associations between various neuromarkers and treatment response but did not use an out-of-sample validation to assess their predictive power are not discussed (for these studies, see recent reviews: Dichter, Gibbs, & Smoski, 2015; Fonseka et al., 2018; Kang & Cho, 2020; Olbrich & Arns, 2013; Phillips et al., 2015; Taylor et al., 2021). Additionally, in the interest of brevity, studies with sample sizes of 30 and below will not be discussed because of their limited generalizability (see Tables 1 and 2 for study summaries).

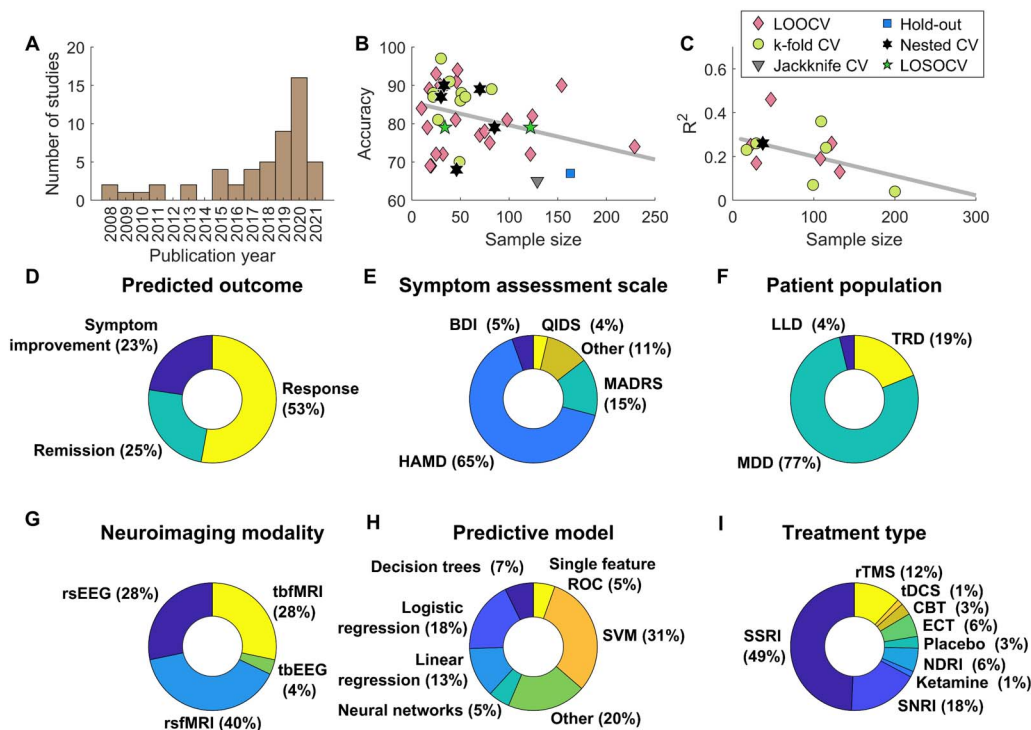
The literature search was performed during the period of June to October 2021. Using the search terms *depress\* AND treatment response prediction AND (accuracy OR sensitivity OR specificity OR regression) AND (fMRI OR EEG)* on PubMed yielded 328 articles. Following a double rater assessment procedure, 35 articles met the inclusion criteria, with the main reasons for exclusion being (a) reviews and meta-analyses; (b) study protocols; (c) studies investigating other disorders; and (d) studies using no validation techniques. A further literature search was conducted using additional sources: (a) reference lists of already qualifying papers and related reviews, and (b) a search engine ResearchRabbit ([www.researchrabbit.ai](http://www.researchrabbit.ai)), which allows for the discovery of the most related papers based on a collection of input papers. This revealed 18 additional articles, leading to a total of 53 articles, which are reviewed below (Figure 2, Tables 1 and 2).

### Strongly Data-Driven Studies

**rsEEG.** Resting-state EEG (rsEEG) studies are used to evaluate intrinsic neural activity, which is not elicited through a specific task. Resting-state measures often require less domain-specific knowledge, while EEG itself is inexpensive and quick to administer. EEG signals are often classified into frequency bands (i.e., delta, theta, alpha, beta, and gamma), each of which have been associated with different brain states (e.g., sleep, rest, alertness). Common EEG measures used as features for treatment response prediction include absolute and relative power, as well as coherence of frequency bands, and the majority of data-driven rsEEG studies have focused on predicting first-line antidepressant response (i.e., selective serotonin reuptake inhibitors, SSRIs).

While features from all frequency bands may be valuable for predicting first-line antidepressant response, the most predictive features tend to be from either alpha and/or theta bands. A recent study by Jaworska, de la Salle, Ibrahim, Blier, and Knott (2019) found alpha and theta power in the frontoparietal area to be highly predictive of SSRI (escitalopram) and norepinephrine–dopamine reuptake inhibitor (NDRI) (bupropion) response, with clinical, EEG, and current source density measures achieving 88% accuracy using an SVM classifier. Zhdanov et al. (2020) demonstrated similar results, whereby high alpha band power in the ACC was highly predictive of escitalopram response. An SVM classifier trained on clinical and EEG features with leave-one-site-out cross-validation (LOSOCV) yielded a slightly lower balanced accuracy (BAC) of 79%.

W. Wu et al. (2020) and Rajpurkar et al. (2020) analyzed rsEEG data from two large multisite clinical trials for depression: EMBARC (Establishing Moderators and Biosignatures of Antidepressant Response in Clinic Care), and iSPOT-D (International Study to Predict Optimized Treatment for Depression), respectively. W. Wu et al. (2020) introduced their Sparse EEG Latent SpacE Regression algorithm to predict sertraline response, and found that only alpha band power—and not theta, beta, or gamma—significantly predicted symptom improvement ( $r = 0.60$ ,  $p = 2.88 \times 10^{-11}$ ). The same model could not predict symptom improvement in the placebo group, suggesting that the model captured features unique to the antidepressant



**Figure 2.** A visual summary of the reviewed studies. (A) The number of studies published each year. (B) Reported outcome prediction accuracy and (C) coefficient of determination as a function of sample size and broken down by validation methods used. LOOCV: leave-one-out cross-validation. LOSOCV: leave-one-site-out cross-validation. (C–I) Statistics of the reviewed 53 studies on the relevant metrics of study design. (D) The definition of outcomes. Symptom improvement: predicting symptom improvement after treatment on a continuous scale. Remission: predicting whether a certain threshold (e.g.,  $\leq 7$  total score on the 17-item Hamilton Rating Scale for Depression, HAMD-17, after treatment) will be reached. Response: predicting whether a certain amount of reduction in symptoms will be reached (e.g.,  $\geq 50\%$  reduction in the total HAMD score after treatment). (E) The scale used to assess depressive symptoms. (F) The studied patient population. MDD: major depressive disorder. TRD: treatment-resistant depression. LLD: late-life depression. (G) The neuroimaging modality used for data collection. rs: resting state. tb: task-based. (H) The model class used for treatment response prediction. ROC: receiver operating characteristic. (I) The treatment type for which predictions were made. SSRI: selective serotonin reuptake inhibitor. SNRI: serotonin-norepinephrine reuptake inhibitor. NDRI: norepinephrine-dopamine reuptake inhibitor. tDCS: transcranial direct-current stimulation. rTMS: repetitive transcranial magnetic stimulation. ECT: electroconvulsive therapy. CBT: cognitive behavioral therapy.

response. The authors further validated the model on an independent rsEEG dataset from Fonzo et al. (2019), which yielded a similar classification performance based on fMRI recordings ( $r = 0.44$ ,  $p = 0.02$ ; see the section *Theory-Driven Task Design* for discussion on the study), thus providing evidence for a neurobiological phenotype that can be detected across neuroimaging modalities. Rajpurkar et al. (2020) predicted *individual* symptom improvement in response to three SSRIs (escitalopram, sertraline, or venlafaxine). Baseline symptom scores and relevant EEG features were used to train a gradient-boosted decision tree model achieving a concordance index of  $\geq 0.8$  on 12 out of 21 clinician-rated symptoms ( $R^2$  0.3–0.7). Concordance index indicates the probability that, given two random patients, the algorithm will correctly identify which patient showed greater improvement. Surprisingly, the inclusion of treatment groups did not significantly improve model performance, suggesting that the EEG markers were general predictors of treatment outcome instead of predictors of differential treatment response.

For patients who fail to respond to first-line antidepressants, or who are unable to tolerate medications, repetitive transcranial magnetic stimulation (rTMS) may be used as an alternative

treatment. N. Bailey et al. (2019) used primarily alpha and theta frequency band features to predict rTMS response, achieving 87% accuracy, and found that rTMS responders showed elevated theta connectivity and lower alpha power at baseline and Week 1, while nonresponders showed typical theta connectivity (similar to that of controls). However, the same group later attempted to replicate these results using a large independent dataset ( $N = 193$ ), and found no significant difference in theta connectivity or alpha power between rTMS responders and nonresponders (N. W. Bailey et al., 2021). Hasanzadeh, Mohebbi, and Rostami (2019) used a combination of power spectrum features across five frequency bands to predict rTMS response in MDD, achieving 91% BAC, and found that rTMS responders had significantly lower baseline beta power.

**rsfMRI.** rsfMRI measures spontaneous brain activity believed to reflect functional communication between spatially distributed brain regions (van den Heuvel & Pol, 2010). Growing literature supports the notion that depression is associated with widespread aberrant functional connectivity, mainly in frontostriatal and limbic brain networks (Greicius et al., 2007; Mayberg et al., 2005; Sheline et al., 2009). Hence resting-state functional connectivity (rsFC) patterns, either in specific networks or on a whole-brain scale, can provide a platform for investigating the hypothesized functional disconnectivity effects in MDD and network changes in response to treatment.

Similar to rsEEG, a majority of rsfMRI studies investigated the impact of first-line pharmacotherapeutic interventions, and hyperconnectivity in the default mode network (DMN) was a common emergent finding across studies. A multisite study by Tian et al. (2020) applied a multilayer modularity framework with leave-one-site-out CV to predict SSRI (escitalopram) response using whole-brain features of functional segregation and integration, achieving accuracy rates of 69–72% across sites. Treatment responders showed stronger connections between the ACC and nodes within the DMN, suggesting that high interactions of the ACC with other regions may be predictive of treatment response. Chin Fatt et al. (2020) predicted SSRI (sertraline) versus placebo response using major cortical rs-networks and subcortical regions with a linear mixed model ( $r = 0.22$ – $0.36$ ). In general, the authors found that higher connectivity within the DMN and between the DMN and the executive control network predicted better treatment outcome. Using data from the international iSPOT-D trial, Korgaonkar, Goldstein-Piekarski, Fornito, and Williams (2020) predicted SSRI (escitalopram or sertraline) and SNRI (venlafaxine) response using intrinsic FC derived from task-based fMRI data. The authors found that irrespective of medication type, greater connectivity within the DMN was predictive of treatment remission (BAC 68%). In comparison, using a functional connectome “fingerprint” at one week posttreatment, Nemati et al. (2020) found that, compared with placebo, reductions in the DMN predicted better response to sertraline ( $r = 0.27$ ), suggesting an early pattern of normalization in the DMN. Furthermore, the authors investigated the generalizability of the sertraline connectome fingerprint in predicting ketamine (rapid antidepressant) response. The model predicted ketamine response compared with an active control (lanicemine;  $r = 0.57$ ), but failed to predict ketamine response compared with placebo ( $p > 0.05$ ). Using the same study sample, Fan et al. (2020) identified a baseline functional connectome fingerprint that significantly predicted symptom improvement irrespective of treatment type (sertraline or placebo), but unlike Nemati et al. (2020), failed to predict the antidepressant treatment-specific response.

A couple of studies used whole-brain rsFC to investigate treatment response in mixed-treatment MDD cohorts (medications included SSRIs, SNRIs, and NDRIs). Ju et al. (2020) used a linear regression model and significantly predicted symptom improvement ( $r = 0.43$ ,

$p = 2.73 \times 10^{-6}$ ) at one month, and in out-of-sample patients for up to three months. Kong et al. (2021) developed a novel spatiotemporal graph convolutional network (STGCN) framework, which predicted treatment response with 90% accuracy.

A study by Drysdale et al. (2017) used whole-brain resting-state networks to investigate differential rTMS response between four depressive subtypes, which are neurophysiological subtypes defined by distinct patterns of dysfunctional connectivity in limbic and frontostriatal networks. Using connectivity features and this biotype classification, an SVM classifier was able to predict rTMS response with 90% accuracy. The most discriminating connectivity features involved the dorsomedial prefrontal (dmPFC) stimulation target and the left amygdala. The final model was further validated on an independent replication set ( $n = 30$ ) and obtained comparable accuracy rates (88–93%). However, when Dinga et al. (2019) attempted to replicate depressive subtypes identified by Drysdale and colleagues on an independent sample ( $n = 187$ ), the authors were unable to replicate these findings and found the methodology to be unreliable in their sample.

Finally, several studies have investigated rsFC patterns associated with electroconvulsive therapy (ECT) response. Van Waarde et al. (2015) used standard group ICA to extract rs-networks, each of which trained an SVM classifier, and found that rs-networks centered in the dmPFC (BAC = 85%) and ACC (BAC = 78%) significantly predicted ECT response in severe and treatment-resistant depression. Leaver et al. (2018) used multimodal fMRI metrics to predict ECT response achieving significant BACs (58–68%). Notably, the left dlPFC and subgenual ACC, both targets of rTMS, as well as connectivity between motor and temporal networks (near ECT electrodes), were consistently identified as informative features in the models. Finally, a larger study ( $n = 122$ ) by Sun et al. (2020) used whole-brain rsFC to train a connectome-based model to predict depressive rating changes and remission status following ECT. Negative FC networks (anti-correlated with changes in depressive scores) were the most predictive ( $r = 0.51$ , accuracy = 76%), with FC between the inferior frontal gyrus and temporal regions demonstrating the most predictive power.

### **Combination of Data-Driven and Theory-Driven Methods**

Up until now we have discussed studies that relied solely on data-driven approaches throughout the analysis pipeline. In this section, we will discuss studies that combined data-driven and theory-driven methods by considering two dimensions: data collection and dimensionality reduction.

**Theory-driven task design.** Several studies have incorporated domain knowledge at the level of the study design itself, that is, by measuring brain activity during cognitive tasks that probe specific mechanisms previously shown to be implicated in MDD. Broadly, these mechanisms relate to reward processing and emotion regulation (Han et al., 2021; Phillips et al., 2015; Stuhmann, Suslow, & Dannlowski, 2011).

**Task-Based EEG.** Using an EEG visual oddball task to predict SSRI response, Mumtaz, Xia, Mohd Yasin, Azhar Ali, and Malik (2017) compared three time-frequency decomposition techniques for feature extraction. EEG-based wavelet features extracted from frontal and temporal areas involving delta and theta frequency bands were the most predictive, with a logistic regression model producing an accuracy of 92%. N. Bailey et al. (2018) employed an EEG working memory task to evaluate rTMS response in patients with treatment-resistant depression achieving 91% BAC with an SVM classifier. At baseline and Week 1, responders showed enhanced fronto-midline theta power and higher theta connectivity compared with

nonresponders. Although, using rsEEG data, increased theta connectivity in rTMS responders was later replicated by the same group (N. Bailey et al., 2019), this finding was subsequently disproved using a large independent dataset (N. W. Bailey et al., 2021).

**Task-Based fMRI.** The majority of task-based studies employed emotional paradigms to probe abnormal processing of emotional stimuli commonly implicated in MDD (Stuhrmann et al., 2011). Using an emotional faces task, Godlewska et al. (2018) predicted SSRI (escitalopram) response based on the mean ACC activity to sad versus happy facial expressions. A moderate accuracy of 72% was achieved with responders showing increased pretreatment pregenual ACC activity to sad versus happy faces. Fonzo et al. (2019) used an emotional conflict task to predict symptom improvement following SSRI (sertraline) treatment and found that a greater downregulation of conflict-responsive regions predicted better outcomes ( $r = 0.49$ ,  $p < 0.001$ ). The same model could not predict improvement in the placebo group, suggesting that the model captured features unique to the antidepressant. Karim et al. (2018) employed an emotion regulation and emotion reactivity task, as well as rsfMRI to predict SNRI (venlafaxine) response in late-life depression (LLD). Using whole-brain connectivity and regions of task activation at baseline and one day following treatment, a logistic regression model achieved an AUC (area under the curve) of 77%, outperforming the use of baseline fMRI alone. The majority of predictive regions were in the frontal cortex, with the emotional reactivity task producing the most informative features.

Two studies assessed patients from the iSPOT-D trial (SSRI/SNRI treatment), whereby an emotional faces paradigm was used to investigate amygdala reactivity. Williams et al. (2015) found that amygdala hyporeactivity to subliminal happy and threat expressions was a general predictor of treatment response (accuracy 75%). However, amygdala reactivity to subliminal sadness functioned as a differential predictor, whereby nonresponders to SNRIs showed pretreatment hyperreactivity to subliminal sadness, which progressed to hyporeactivity posttreatment, and predicted SNRI response with 77% accuracy. Goldstein-Piekarski et al. (2016) investigated the interaction between amygdala engagement and early life stress to predict “functional remission,” which the authors defined by combining measures of the clinician-rated HAMD, self-reported 16-item Quick Inventory of Depressive Symptomatology–Self-Rated (QIDS-SR16), and observer-rated functional capacity using the Social and Occupational Functioning Assessment Scale. A discriminant analysis yielded a BAC of 77%, and similar to Williams et al. (2015), in patients with low early life stress, lower amygdala reactivity to both happy and threat-related stimuli increased the likelihood of remission. In comparison, for those with high exposure to early life stress, greater amygdala reactivity to happy stimuli predicted remission.

Using a Go/No-Go task, Tozzi, Goldstein-Piekarski, Korgaonkar, and Williams (2020) investigated differential response to SSRIs (escitalopram and sertraline) and SNRIs (venlafaxine). Connectivity between the dlPFC and the supramarginal gyrus (SMG) and between SMG and the middle temporal gyrus (MTG) was associated with response to sertraline and venlafaxine, but not to escitalopram. Using baseline symptom scores and the mean FC contrast values as inputs to a logistic regression classifier, venlafaxine response was predicted with BAC of 79%, while sertraline response was predicted with BAC of 84%. Interestingly, higher FC between both dlPFC-SMG and SMG-MTG was associated with response to sertraline, whereas lower connectivity was associated with response to venlafaxine.

Nguyen et al. (2019) employed a reward processing task and deep learning model to predict depressive rating changes and response status following NDRI (bupropion) treatment. Regions of activation were extracted from two contrast maps, one for anticipation and the

other for reward expectation. The final model achieved a root mean square error of 4.71 ( $R^2 = 0.26$ , AUC = 71%), and important clusters included the medial frontal cortex, amygdala, cingulate cortex, and striatum. The final model performed poorly (negative  $R^2$ ) on SSRI (sertraline) and placebo-treated subjects from the same dataset, suggesting that the model likely learned features specific to the bupropion response. Finally, a recent study by Brandt et al. (2021) used a comparatively large MDD cohort ( $n = 90$ ) and found that pretreatment reward-related brain activity was not predictive of SSRI (escitalopram) treatment response. The authors also found no differences in reward reactivity estimates between healthy and depressed individuals and no change following eight weeks of treatment.

### Theory-driven feature selection

**rsEEG.** Two studies used quantitative EEG and theory-driven feature selection, particularly measures in the theta band over frontal regions, to predict treatment outcome. Cook, Hunter, Caudill, Abrams, and Leuchter (2020) predicted SSRI (escitalopram) and NDRI (bupropion) remission using a previously validated biomarker, the Antidepressant Treatment Response (ATR) index (Cook et al., 2020; Leuchter et al., 2009), which combines theta and alpha power metrics at baseline and Week 1 in frontotemporal channels. Higher ATR values were predictive of SSRI remission (BAC 65%), but not NDRI remission, which was selected as a control comparison. Erguzel et al. (2015) used an artificial neural network to predict rTMS response in patients concurrently receiving SSRIs, and using frontal cordance values from theta and delta bands, the model achieved BACs of 86–89%.

**rsfMRI.** Based on previous literature, two studies selected a subset of resting-state features for the prediction of SSRI and SNRI response. In late-life depression, Patel et al. (2015) employed an Alternating Decision Tree model to predict antidepressant response yielding 89% accuracy based on cognitive scores, structural, and rsFC features (in the default mode network, DMN, and anterior salience network, aSN). Fewer structural connections in the aSN was predictive of response, while lower FC in the dorsal DMN was predictive of nonresponse. Pei et al. (2020) used an SVM classifier based on the FC of 14 priori brain regions and predicted antidepressant response with a BAC of 81%. Notably, a model trained using whole-brain features achieved the same accuracy (81%).

Similar to Korgaonkar et al. (2020), Braund et al. (2022) used iSPOT-D data (SSRI/SNRI treatment) to investigate intrinsic FC networks that characterized neuroticism in 229 MDD patients, and using an SVM predicted treatment response with a BAC of 74%. Greater connectivity within and between the salience, executive control, and somatomotor brain networks was associated with higher baseline neuroticism. Irrespective of treatment type, increased network activity was predictive of poorer treatment outcomes that was not mediated by baseline neuroticism. Goldstein-Piekarski et al. (2018) also used iSPOT-D data, to predict treatment remission using connectivity within the DMN, with a focus on the posterior cingulate cortex (PCC). Connectivity between the PCC and ACC/mPFC together was predictive of remission (82% accuracy), whereby nonremitters showed relative hypoconnectivity compared with remitters, who showed intact connectivity similar to that of controls. Differential prediction of remission using PCC connectivity did not survive correction.

Cash et al. (2019) significantly predicted rTMS response using rsFC in the DMN and affective network (AN), as well as BOLD signal power and Week 1 clinical response ( $r = 0.68$ ,  $p < 0.001$ , 93% accuracy). However, some significant relationships between individual features and treatment outcome were only observed once participants with the lowest treatment outcome (<0% change) were omitted. Hopman et al. (2021) attempted to replicate previous

findings that stronger dlPFC-sgACC anticorrelated connectivity was associated with rTMS response; however, the authors could not confirm this relationship. Instead, using seed-based features of the left dlPFC (rTMS target) and sgACC, the authors predicted rTMS response with 89% accuracy, finding that greater connectivity disruptions involving the central executive network was associated with poorer response.

## GENERATIVE MODELING AND GENERATIVE EMBEDDING

Compared with the studies reviewed so far, generative embedding approaches have the potential to better capture disease mechanisms and provide more interpretable treatment response predictions (Brodersen et al., 2011; Frässle et al., 2018). Generative embedding incorporates generative models of information processing dynamics in the brain. By fitting these models to (brain or behavioral) data, one can effectively reduce the dimensionality of the raw data to a handful of highly informative model parameter estimates—that is, mechanistically interpretable features. These parameter estimates are then used as input features in machine learning algorithms to predict treatment response. Depending on how well these models capture mechanisms relevant for treatment response prediction, this approach could improve not only the interpretability of the predictions but also their accuracy. However, generative embedding is yet to be fully utilized in the context of treatment response prediction in MDD. In this section, we would like to highlight several lines of research employing generative modeling approaches that have been used to study MDD but have not been applied to treatment response prediction.

### *Generative Models of Brain Data*

One way to model the information processing dynamics in the brain is by using dynamic causal modeling (DCM), which allows for estimation of directed interdependencies (i.e., effective connectivity) among multiple brain regions (Friston et al., 2019; Stephan & Friston, 2010). Unlike functional connectivity, which describes temporal correlations in BOLD responses across brain regions, effective connectivity rests on a generative model, which specifies directed relationships between populations of neurons (Friston, 2011). DCM is thus able to account for asymmetries in forward and backward connections (Bastos et al., 2012; Felleman & Van Essen, 1991; Markov et al., 2014), which creates a possibility for a more detailed characterization of disease mechanisms. Furthermore, recent studies using DCM for electrophysiological data were able to incorporate microscale details such as the conductance of specific receptor populations (Gilbert et al., 2016; Moran, Symmonds, Stephan, Friston, & Dolan, 2011; Schöbi et al., 2021; Symmonds et al., 2018), demonstrating the potential of DCM for multilevel description (from micro- to macroscale) of disease mechanisms. Estimating receptor densities in vivo from noninvasive EEG recordings would be a major step towards linking psychiatric symptoms to the mechanisms of action of pharmacological interventions that target specific neurotransmitter systems.

While many studies have applied DCM for studying mechanisms underlying MDD (Li et al., 2018), so far very few have used it for treatment response prediction. Vai et al. (2016) used effective connectivity measures obtained from DCM to investigate treatment response to escitalopram. Pretreatment effective connectivity during emotional face processing was found to discriminate nonremitters from remitters and controls after six weeks of treatment. Nonremitters showed reduced endogenous connectivity from the amygdala to the ventrolateral PFC and to the ACC, and increased modulation of the ACC to the amygdala when processing fearful

faces. However, these results were obtained by performing a series of *t* tests and there was no analysis of how accurately these effects could predict treatment at the individual level.

Frässle et al. (2020) is the only study to apply a generative embedding approach for predicting illness course in MDD. Unlike the studies reviewed so far, this study aimed to predict long-term (up to two years) trajectories of MDD in a naturalistic cohort (participants were receiving mixed treatments), and thus it is not directly comparable, but it serves to illustrate how generative embedding could be applied for treatment response prediction. In this study, illness course was defined by dividing participants into three different groups: remitted (rapid remission); improved (slow remission); and chronic (treatment-resistant). Pretreatment fMRI data recorded during an emotional face perception task was used for predicting illness course. Six regions of interest (ROIs) were selected based on their association in the literature with the extended face perception network: bilateral occipital face area, fusiform face area, and amygdala. The best model allowed emotion processing to modulate forward and backward intra- and interhemispheric connections among homotopic brain regions. Using effective connectivity parameters as features, SVM predicted chronic versus remitted groups with BAC of 79%, and improved versus remitted groups with BAC of 61%. Importantly, this performance exceeded that of conventional non-generative methods that used functional connectivity or local activation (computed from the same network of ROIs) as features for classification; these did not result in above-chance performance.

### **Generative Models of Behavioral Data**

While generative models of brain dynamics can help us explain neural data, they fall short of explicitly linking these dynamics to behavior. Given that relevant clinical symptoms manifest in behavior, explaining observable behavior is an important consideration for generative models. In the research setting, MDD has been associated with deficits in value-based decision-making, especially in tasks involving reinforcement learning and expectations about the future (Eshel & Roiser, 2010; Mukherjee, Lee, Kazinka, Satterthwaite, & Kable, 2020; Must, Horvath, Nemeth, & Janka, 2013). Not surprisingly, one of the most popular frameworks for modeling these aspects of behavior in MDD has been reinforcement learning (RL), which models adaptive decision-making in the face of rewards and punishments (Chen, Takahashi, Nakagawa, Inoue, & Kusumi, 2015; Huys, Pizzagalli, Bogdan, & Dayan, 2013). The central variable in RL is the reward prediction error (RPE)—the difference between expected and observed reward/punishment—which guides learning of value of different stimuli, which in turn guides actions. Importantly, RL modeling approaches make it possible to study reward processing in a lot more detail and investigate how different elements of the decision-making process such as RPE, expected value (“wanting”), reward sensitivity (“liking”), learning rate, memory of previous reinforcement, noisiness of action selection, and so on, might be implicated in MDD and how they might relate to different MDD subtypes (Chen et al., 2015; Robinson & Chase, 2017; Ruppel, Stankevicius, Huys, Steele, & Seriès, 2018). Crucially, even though RL models are fitted to behavioral data, the computational processes described by them can then be used to investigate brain activity specific to this process by including model variables as regressors in a general linear model (GLM), which in neuroimaging data analysis, has been known as **model-based fMRI** (Gläscher, Daw, Dayan, & O’Doherty, 2010; Katahira & Toyama, 2021; O’Doherty, Hampton, & Kim, 2007).

To date, only one study has applied RL model-based fMRI for treatment response prediction in MDD. Queirazza, Fouragnan, Steele, Cavanagh, and Philiastides (2019) aimed to obtain

Model-based fMRI:  
Extracts features of latent cognitive mechanisms using theory-driven generative models of behavioral data and uses those features as regressors in GLM-based fMRI analysis to find neural correlates.



mechanistically meaningful fMRI predictors of response to computerized cognitive behavioral therapy (cCBT) based on pretreatment brain activity during a probabilistic win/loss reversal-learning task. The data were analyzed by first fitting an RL model to the behavioral data to estimate trial-wise RPEs. Next, RPEs, weighted by a dynamic learning rate, were used as regressors in a GLM when analyzing the fMRI data. Finally, the resulting regression coefficients were used to predict cCBT treatment response using an SVM, RVM, and logistic regression. All classifiers showed comparable performance, with the SVM yielding the best performance (BAC 72%). Neural activity encoding-weighted RPEs in the right striatum and right amygdala were the most discriminative features of treatment response, with greater pretreatment activity predicting better response. The authors suggested that greater neural signaling of the weighted RPE might make cognitive restructuring practiced during cCBT more effective, fostering more balanced beliefs about the self and the world.

MDD can also be understood within a more general computational framework of Bayesian inference (Badcock, Davey, Whittle, Allen, & Friston, 2017; Barrett, Quigley, & Hamilton, 2016; Chekroud, 2015; Huys, Daw, & Dayan, 2015; Kube, Schwarting, Rozenkrantz, Glombiewski, & Rief, 2020; Paulus & Angela, 2012; Stephan et al., 2016). Central to Bayesian accounts of decision-making is the observation that external (world) states and internal (bodily) states cannot be directly observed and must be inferred from ambiguous sensory information. Similarly, different action policies and their consequences also carry varying degrees of uncertainty. All these types of uncertainty shape one's expectations of rewarding or aversive outcomes (cf. reinforcement learning), but even more importantly, they determine how information is sampled (via action selection and attention) and to what extent the resulting positive or negative experiences are integrated into one's model of the world. Within this framework, deficits in value-based decision-making in MDD can be cast as a biased construction of internal and external states, which results in a maladaptive positive feedback loop involving one's model of the world, action selection, and mood. The biases themselves could result from a miscalibration of precision (the inverse of uncertainty) associated with prediction errors (PEs)—which, similarly to RPEs, represent the difference between expected and experienced sensory input and guide learning (Badcock et al., 2017; Barrett et al., 2016; Kube et al., 2020). For instance, attenuation of PEs would make one immune to changing one's negative beliefs in the face of disconfirming evidence (Barrett et al., 2016; Kube et al., 2020). On the other hand, increased precision of PEs for social contexts may increase sensitivity and attention to interpersonal cues and could lead to social withdrawal and anhedonia (Badcock et al., 2017). Finally, in the context of interoception, a miscalibration of the precision associated with PEs would lead to a disruption of allostasis (i.e., the brain's ability to anticipate and flexibly adapt to changing metabolic needs), which could explain fatigue and inflammation observed in MDD (Barrett et al., 2016; Stephan et al., 2016). Relevantly, recent preliminary findings by Hough et al. (2021) indicate that pretreatment allostatic load and metabolic dysregulation might be predictive of SSRI response. Thus, adopting a Bayesian framework allows for integration of brain and behavior as well as the body and physiological states. This potentially affords a much more comprehensive picture of MDD. So far, to the best of our knowledge, no studies have applied these models for treatment response prediction.

### **OTHER REMAINING CHALLENGES**

While theory-driven generative models have the potential to improve our understanding of MDD as well as capture interpretable and discriminative features for treatment response prediction, multiple other challenges remain.

### **Treatment Response Prediction Versus Differential Treatment Response Prediction**

Most of the studies reviewed in this paper have focused on building a model that can predict treatment response for a single intervention. In clinical practice, a more useful prediction would be *differential* treatment response prediction, indicating which of *several* available treatments is the most likely to lead to improvement for a given patient (Dunlop & Mayberg, 2014; Perlman et al., 2019). This of course could be determined by applying individual models for each treatment and then aggregating the results to derive differential prediction. However, if different models have to rely on different data modalities and require separate validations, this would be much more time- and resource-intensive. Therefore, a single model that can reliably predict treatment response to multiple treatment options would be much more useful in practice.

Several of the reviewed studies did involve multiple treatments, but differential treatment response prediction remains very limited. A tbfMRI study by Tozzi et al. (2020) found that greater functional connectivity between both dlPFC-SMG and SMG-MTG during behavioral inhibition was associated with response to SSRI sertraline, whereas lower connectivity was associated with response to SNRI venlafaxine; however, their model failed to predict response to SSRI escitalopram. Another tbfMRI study by Williams et al. (2015) found amygdala hyporeactivity to subliminal happy and threat expressions to predict treatment response to escitalopram, sertraline, and venlafaxine, while amygdala hyperreactivity to sad expressions specifically predicted nonresponse to venlafaxine. Investigating the same three treatments, two rsfMRI (Goldstein-Piekarski et al., 2018; Korgaonkar et al., 2020) and one rsEEG (Rajpurkar et al., 2020) studies, were only able to predict treatment response across the treatments but not differential treatment response. Some studies involving multiple treatments did not investigate differential treatment response despite having large enough sample sizes (Braund et al., 2022; Cook et al., 2020; Pei et al., 2020), while others were limited by too small samples (Crane et al., 2017; Jaworska et al., 2019; Ju et al., 2020; Khodayari-Rostamabad, Reilly, Hasey, de Bruin, & MacCrimmon, 2013; Patel et al., 2015; Rabinoff, Kitchen, Cook, & Leuchter, 2011; Sikora et al., 2016).

For more clinically relevant results, future studies should focus on establishing predictions of differential treatment response and include a wider range of treatments (e-psychotherapy, ECT, rTMS, first-line antidepressants, rapid-acting antidepressants, psychedelic therapy). Bearing in mind the heterogeneity of MDD, this would be of more benefit than comparing very few treatments with similar mechanisms of action (e.g., escitalopram vs. sertraline). This, however, would require large samples and multisite studies, which can be practically challenging. One of the largest current projects, iSPOT-D (Williams et al., 2011), aims to recruit  $N = 672$  per treatment arm but includes only three treatments (escitalopram, sertraline, and venlafaxine).

### **Defining Outcomes**

**Problems with dichotomization.** The majority of studies reviewed here focused on predicting either treatment response or remission by dichotomizing the reduction in symptoms of depression—with *response* often defined as  $\geq 50\%$  reduction in symptoms from the baseline and *remission* defined as reaching a score below some low threshold, such as  $\leq 7$  for HAMD-17. This is rather problematic because dichotomizing continuous variables leads to a loss of information and thus a loss of statistical power (Altman & Royston, 2006; Maxwell & Delaney, 1993). For example, a patient with 49% reduction in symptoms will be considered to be the same as a patient with 0% reduction (both being “nonresponders”) but categorically different from a patient with 50% reduction in symptoms (a “responder”). Ignoring this within-group

variability could lead to false inferences about the features that are predictive of treatment response (Austin & Brunner, 2004). Even more importantly, such coarse categorization undermines the very possibility of accurately predicting treatment outcomes.

Thirteen studies have circumvented this problem by applying regression models for predicting the symptom improvement on a continuous scale (see Tables 1 and 2), and one study (Rajpurkar et al., 2020) predicted improvement in individual symptoms, rather than summed symptom scores. Although the majority of studies produced strong results (median  $R^2 = 0.24$ ), it is not straightforward to compare them with the classification accuracy reported in the other studies that dichotomized symptom scores. In principle, predicted symptom scores could be converted to classification accuracy by post hoc dichotomization. This was exemplified by two studies, whereby the model's regression outputs were dichotomized post hoc to obtain remission (Nguyen et al., 2019; Sun et al., 2020) prediction accuracy. Converting symptom improvement predictions into classification accuracy might also have benefits in clinical applications.

Another consideration when modeling symptom improvement is whether baseline depression severity should be included as a predictor or used to derive the predicted outcome (the difference between baseline and posttreatment depression severity). In some cases, these options can lead to diverging results (Lord, 1967; Tu, Gunnell, & Gilthorpe, 2008). However, when the baseline exhibits collinearity with other predictors of interest—as we might expect to be the case for depression severity and brain activity—including baseline as a predictor has been shown to lead to biased results (Castro-Schilo & Grimm, 2018; Farmus, Arpin-Cribbie, & Cribbie, 2019; Liu, Lu, Mogg, Mallick, & Mehrotra, 2009). In line with this, all symptom improvement studies reviewed here did not include baseline severity as a predictor. Note that when predicting dichotomized outcomes, baseline severity is also implicitly included in the outcome definition for the classification of responders versus nonresponders, because treatment response is usually defined as percentage change scores (e.g., 50% symptom improvement compared with baseline). Conversely, in the case of remitters versus nonremitters, the baseline does not feature at all as remission is defined as absence of clinically relevant symptoms at follow-up regardless of baseline symptom severity. However, among the reviewed studies that predicted dichotomized outcomes, the vast majority tested and reported no differences in baseline scores between the groups, which makes the considerations of baseline effects not relevant for the interpretation of the results. When baseline differences do exist in the context of predicting remission, including baseline among predictors should be considered (e.g., Karim et al., 2018).

**Predicting improvement of individual symptoms to address MDD heterogeneity.** It is also worth considering whether using the overall improvement across symptoms is sufficiently informative for predicting treatment outcomes. MDD is a broad category, and diagnosed individuals display diverse symptom profiles, with some sharing no symptoms in common (E. I. Fried, 2017; E. I. Fried, Nesse, Zivin, Guille, & Sen, 2014; Goldberg, 2011). This heterogeneity is also reflected in the most commonly used MDD rating scales, such as HAMD (Hamilton, 1967), MADRS (Montgomery & Åsberg, 1979), BDI (Schwab, Bialow, Clemmons, Martin, & Holzer, 1967), and QIDS (Rush et al., 2003); these instruments differ considerably in their item content because they place different emphasis on different MDD symptoms (E. I. Fried, 2017). Developing models that can predict treatment effects on each individual symptom, thus, would make it possible to account for heterogeneity in MDD and would prevent problems of comparing study findings based on total depression scores that are derived from instruments that overlap but that are not interchangeable. One exemplar study reviewed here (Rajpurkar et al., 2020) addressed this issue by building a model to predict the change in each of the 21 items in

HAMD-21 based on rsEEG data. The authors used gradient-boosted decision trees (GBDT), which allows for modeling of nonlinear associations. Their model achieved a concordance index of  $\geq 0.8$  on 12 out of the 21 symptoms ( $R^2$  of 0.32–0.7). These results are encouraging and suggest that prediction of treatment effects on individual MDD symptoms is feasible.

**Patient-centered definition of outcomes: Functional recovery and quality of life.** Another consideration is that the widely adopted depression scales used to assess symptom severity do not capture all relevant aspects of treatment outcomes. From a patient's point of view, symptom resolution is only one factor in determining remission from depression. Positive aspects of mental health (Zimmerman et al., 2006), functional recovery (Greer, Kurian, & Trivedi, 2010; Lam, Parikh, Michalak, Dewa, & Kennedy, 2015; Oluboka et al., 2018), and quality of life (QoL; IsHak et al., 2011) are just as important. MDD-related functional impairment can span multiple domains such as occupational, social, physical, and cognitive functioning, for which both objective and subjective measures exist. QoL highly overlaps with functional impairment measures but often assesses well-being across a wider variety of domains, including emotional well-being and overall life satisfaction, and is often based on subjective self-rating of these domains (Endicott, Nee, Harrison, & Blumenthal, 1993; Frisch et al., 2005; Leon et al., 1999). The impact of MDD on the aforementioned domains is not always well captured by the scales designed to assess depressive symptoms, while remission of depressive symptoms alone does not necessarily lead to full functional recovery (Habert et al., 2016) and does not fully restore QoL (IsHak et al., 2013; Morton et al., 2021). Relatedly, the narrow assessment of MDD symptoms does not take into account treatment side effects, which are very common for antidepressant drugs (Read & Williams, 2018). Only one of the reviewed studies (Goldstein-Piekarski et al., 2016) focused on the prediction of functional remission.

Understanding the differential effects of available treatments on these measures is therefore an important research direction. A meta-analysis comparing CBT and SSRI treatments found similar effects on QoL; however, depression improvement was associated with increased QoL only in the CBT group (Hofmann, Curtiss, Carpenter, & Kind, 2017). Another recent study by Fischer et al. (2021) found depression symptom and QoL improvements to be associated with partially distinct changes in functional connectivity of reward neurocircuitry, which was also differentially affected by different antidepressants (sertraline, venlafaxine-XR, and escitalopram). Finally, Koshiyama et al. (2020) found rsEEG beta band power at baseline to be strongly correlated with QoL outcomes at the three-year follow-up, and this correlation was independent of a reduction in symptoms. Following this line of research, to achieve a more complete assessment of treatment outcomes, future studies should consider including functional recovery and QoL among the predicted treatment outcomes.

#### **Reliability of Prediction Accuracy: Sample Size and Validation**

Based on the American Psychiatric Association recommendations, treatment response prediction accuracy of at least 80% would be considered to have clinical utility (Botteron et al., 2012). Most of the reviewed studies report prediction accuracy above 80%, and thus would appear to have sufficient accuracy. However, these numbers are likely to be positively biased. The reviewed studies varied considerably in sample size as well as validation methods (see Tables 1 and 2)—both of which affect the generalizability of the reported prediction accuracy (Gillan & Whelan, 2017; Janssen, Mourão-Miranda, & Schnack, 2018; Kelly et al., 2019; Yarkoni & Westfall, 2017).

A substantial number of studies were not included in this review because of a lack of validation methods. Such studies reported classification accuracy for the data that the classifier

was trained on; this leads to overfitting and poor performance on unseen data (Gillan & Whelan, 2017). To improve generalizability, the majority of studies used internal validation, namely LOOCV or k-fold CV, which trains the classifier on all but a subset of patients and uses the left-out patients for testing model accuracy—repeating this process by permuting all data. However, for the median sample size of the reviewed studies,  $N = 47$ , using LOOCV can still lead to accuracy errors of up to  $\sim 15\%$  (Varoquaux, 2018), with k-fold suffering from similar problems (Vabalas, Gowen, Poliakoff, & Casson, 2019). The least biased method, nested CV, adds an additional layer of validation for testing the model on an unseen portion of the data (Vabalas et al., 2019; Varma & Simon, 2006; Varoquaux et al., 2017). Nested CV is especially important for data-driven feature selection (i.e., filtering) where certain features are removed if they are not associated with the target variable (e.g., treatment response). If done for the whole dataset, this is problematic because it leads to positively biased prediction results. To prevent data leakage, all data preprocessing (e.g. feature selection, imputation, hyperparameter tuning) should be embedded within the cross-validation procedure. Nested CV was used by only six of the reviewed studies (Frässle et al., 2020; Hopman et al., 2021; Khodayari-Rostamabad, Reilly, Hasey, de Bruin, & MacCrimmon, 2010; Leaver et al., 2018; Nguyen et al., 2019; Patel et al., 2015). However, for sample sizes between 100 and 200, which a few reviewed studies had (see Tables 1 and 2), the error associated with the reported accuracy could still be up to 8–10%, regardless of the CV method used (Varoquaux, 2018).

Irrespective of the internal validation used, test cases are randomly selected from the same dataset on which the training is performed. The generalizability of such results can thus be undermined by sample- or site-specific confounds. A good way to protect from site-specific confounds is to perform leave-one-site-out CV (LOSOCV), which was done by two multisite studies: Zhdanov et al. (2020), performing LOSOCV across four different sites ( $N = 122$ ), and Tian et al. (2020), performing LOSOCV across three different sites ( $N = 106$ ). To properly test a model's performance, an independent external dataset, which may differ with respect to clinical assessments (e.g., HAM-D vs. MADRS), inclusion criteria (e.g., as it relates to comorbidities), and data acquisition/preprocessing parameters, is required. For example, the model by N. Bailey et al. (2019), using rsEEG to predict rTMS response in treatment-resistant depression ( $N = 50$ ) and achieving BAC 86% with k-fold CV, was not predictive of treatment response when tested on an independent dataset of  $N = 193$  (N. W. Bailey et al., 2021). Similarly, Drysdale et al. (2017) found four depressive subtypes from rsfMRI data, one of which was associated with large response rates to rTMS treatment; however, these subtypes could not be reproduced in an independent study (Dinga et al., 2019).

The ultimate test of treatment response prediction will require prospective validation studies to assess the actual improvement of outcomes following prediction-based treatment selection as compared with treatment-as-usual (van der Vinne et al., 2021). Such studies will also require careful consideration of how treatment response predictions can be most optimally integrated into the clinical workflow (Kelly et al., 2019).

Two conclusions can be drawn from the generalizability issues discussed above. First, for more informative results, studies should employ validation methods that are less biased, such as nested CV. While more practically challenging, to further increase generalizability of results, studies need to perform external validation of their model predictions. Second, small sample sizes create a bottleneck for any model validation efforts. Thus, multisite studies with larger sample sizes ( $>100$ ) will be crucial for producing sufficiently reliable results. Currently, such initiatives include iSPOT-D (Williams et al., 2011), which aims to recruit  $N = 672$  per treatment arm; Canadian biomarker integration network in depression (CAN-BIND; Kennedy et al., 2012), aiming to recruit  $N = 290$ ; and EMBARC study (Trivedi et al., 2016) with  $N = 160$ .

Larger samples would also help address another issue limiting generalizability: sample bias. The reviewed studies varied in their inclusion criteria, both related to comorbidities (anxiety, bipolar, substance use disorders, etc.) and MDD progression (drug-naïve vs. treatment-resistant vs. late-life depression). Larger samples with more inclusive criteria would be more representative of the broader MDD population, where comorbidities are common (Steffen, Nübel, Jacobi, Bätzing, & Holstiege, 2020).

## CONCLUSION

Because of the heterogeneity in MDD presentation, etiology, and trajectory, the selection of an appropriate treatment by clinicians is a challenging task. This review has detailed the emerging use of computational models for individual treatment response prediction in MDD, highlighting methodological differences along the data-driven and theory-driven spectrum. Although both approaches have shown promising results, multiple challenges remain. Here we argued that a promising research direction for improving interpretability and, potentially, the accuracy of model predictions is theory-driven generative models, which allow for inference on disease mechanisms. Furthermore, we identified several other methodological limitations related to treatment outcome definition and validation of model predictions. The success of translating these tools to clinical practice will depend on carefully designed external validation studies with diverse patient samples and patient-centered outcome measures (see Box 1).

### Box 1. Future directions

Obtaining clinically relevant features:

- Theory-driven generative models of brain activity and behavior can help extract more interpretable and more discriminative features for treatment response prediction.
- A comprehensive model of the mechanisms underlying MDD symptomatology could be built using the hierarchical Bayesian inference framework.

Differential treatment response over treatment response prediction:

- Including more than one and ideally several different treatments would allow for the prediction of *differential* treatment response, which is clinically more useful than predicting response to a single treatment only.

Defining treatment outcomes:

- Predicting symptom improvement on a continuous scale, rather than dichotomized remission or response outcomes, can increase statistical power and prediction accuracy.
- Predicting improvement in individual symptoms can help avoid confounds associated with the diverse symptom profiles of the individuals diagnosed with MDD.
- Going beyond symptom reduction and including functional recovery and quality of life as relevant outcomes can help achieve more patient-centered and thus more relevant predictions of treatment outcomes.

Validating computational models:

- To produce more reliable and generalizable results, more robust validation methods need to be employed. At the very least, nested cross-validation should be used for internal model validation. To further improve generalization, external validation across more than one site is needed.
- Larger sample sizes ( $N > 100$ ) are essential for enabling such validation techniques and ensuring the representativeness of the MDD population.

## AUTHOR CONTRIBUTIONS

Povilas Karvelis: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing – original draft; Writing – review & editing. Colleen E. Charlton: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing – original draft; Writing – review & editing. Shona G. Allohverdi: Conceptualization; Investigation; Writing – original draft; Writing – review & editing. Peter Bedford: Conceptualization; Investigation; Writing – review & editing. Daniel J. Hauke: Conceptualization; Investigation; Writing – review & editing. Andreea Diaconescu: Conceptualization; Funding acquisition; Methodology; Project administration; Supervision; Writing – review & editing.

## FUNDING INFORMATION

Andreea Diaconescu, Krembil Foundation (<https://dx.doi.org/10.13039/501100004089>), Award ID: 1000824. Daniel J. Hauke, Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (<https://dx.doi.org/10.13039/501100001711>), Award ID: 200054.

## REFERENCES

- Adams, R. A., Huys, Q. J., & Roiser, J. P. (2016). Computational psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, *87*(1), 53–63. <https://doi.org/10.1136/jnnp-2015-310737>, PubMed: 26157034
- Al-Kaysi, A. M., Al-Ani, A., Loo, C. K., Powell, T. Y., Martin, D. M., Breakspear, M., & Boonstra, T. W. (2017). Predicting tDCS treatment outcomes of patients with major depressive disorder using automated EEG classification. *Journal of Affective Disorders*, *208*, 597–603. <https://doi.org/10.1016/j.jad.2016.10.021>, PubMed: 28029427
- Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*(7549), 1080. <https://doi.org/10.1136/bmj.332.7549.1080>, PubMed: 16675816
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Arlington, VA: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>
- Austin, P. C., & Brunner, L. J. (2004). Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. *Statistics in Medicine*, *23*(7), 1159–1178. <https://doi.org/10.1002/sim.1687>, PubMed: 15057884
- Badcock, P. B., Davey, C. G., Whittle, S., Allen, N. B., & Friston, K. J. (2017). The depressed brain: An evolutionary systems theory. *Trends in Cognitive Sciences*, *21*(3), 182–194. <https://doi.org/10.1016/j.tics.2017.01.005>, PubMed: 28161288
- Bailey, N., Hoy, K., Rogasch, N., Thomson, R., McQueen, S., Elliot, D., ... Fitzgerald, P. (2018). Responders to rTMS for depression show increased fronto-midline theta and theta connectivity compared to non-responders. *Brain Stimulation*, *11*(1), 190–203. <https://doi.org/10.1016/j.brs.2017.10.015>, PubMed: 29128490
- Bailey, N., Hoy, K., Rogasch, N., Thomson, R., McQueen, S., Elliot, D., ... Fitzgerald, P. (2019). Differentiating responders and non-responders to rTMS treatment for depression after one week using resting EEG connectivity measures. *Journal of Affective Disorders*, *242*, 68–79. <https://doi.org/10.1016/j.jad.2018.08.058>, PubMed: 30172227
- Bailey, N. W., Krepel, N., van Dijk, H., Leuchter, A. F., Vila-Rodriguez, F., Blumberger, D. M., ... Fitzgerald, P. B. (2021). Resting EEG theta connectivity and alpha power to predict repetitive transcranial magnetic stimulation response in depression: A non-replication from the ICON-DB consortium. *Clinical Neurophysiology*, *132*(2), 650–659. <https://doi.org/10.1016/j.clinph.2020.10.018>, PubMed: 33223495
- Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160011. <https://doi.org/10.1098/rstb.2016.0011>, PubMed: 28080969
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711. <https://doi.org/10.1016/j.neuron.2012.10.038>, PubMed: 23177956
- Botteron, K., Carter, C., Castellanos, F. X., Dickstein, D. P., Drevets, W., Kim, K. L., ... Zubieta, J.-K. (2012). *Consensus report of the APA work group on neuroimaging markers of psychiatric disorders*. American Psychiatric Association.
- Brandt, I. M., Köhler-Forsberg, K., Ganz, M., Ozenne, B., Jorgensen, M. B., Poulsen, A., ... Fisher, P. M. (2021). Reward processing in major depressive disorder and prediction of treatment response–neuropharm study. *European Neuropsychopharmacology*, *44*, 23–33. <https://doi.org/10.1016/j.euroneuro.2020.12.010>, PubMed: 33455816
- Braund, T. A., Breukelaar, I. A., Griffiths, K., Tillman, G., Palmer, D. M., Bryant, R., ... Korgaonkar, M. S. (2022). Intrinsic functional connectomes characterize neuroticism in major depressive disorder and predict antidepressant treatment outcomes. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *7*(3), 276–284. <https://doi.org/10.1016/j.bpsc.2021.07.010>, PubMed: 34363999
- Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- 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 Computational Biology*, 7(6), e1002079. <https://doi.org/10.1371/journal.pcbi.1002079>, PubMed: 21731479
- Cash, R. F., Cocchi, L., Anderson, R., Rogachov, A., Kucyi, A., Barnett, A. J., ... Fitzgerald, P. B. (2019). A multivariate neuroimaging biomarker of individual outcome to transcranial magnetic stimulation in depression. *Human Brain Mapping*, 40(16), 4618–4629. <https://doi.org/10.1002/hbm.24725>, PubMed: 31332903
- Castro-Schilo, L., & Grimm, K. J. (2018). Using residualized change versus difference scores for longitudinal research. *Journal of Social and Personal Relationships*, 35(1), 32–58. <https://doi.org/10.1177/0265407517718387>
- Chahal, R., Gotlib, I. H., & Guyer, A. E. (2020). Research review: Brain network connectivity and the heterogeneity of depression in adolescence—A precision mental health perspective. *Journal of Child Psychology and Psychiatry*, 61(12), 1282–1298. <https://doi.org/10.1111/jcpp.13250>, PubMed: 32458453
- Chekroud, A. M. (2015). Unifying treatments for depression: An application of the free energy principle. *Frontiers in Psychology*, 6, 153. <https://doi.org/10.3389/fpsyg.2015.00153>, PubMed: 25750630
- Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience and Biobehavioral Reviews*, 55, 247–267. <https://doi.org/10.1016/j.neubiorev.2015.05.005>, PubMed: 25979140
- Chin Fatt, C. R., Jha, M. K., Cooper, C. M., Fonzo, G., South, C., Grannemann, B., ... Trivedi, M. H. (2020). Effect of intrinsic patterns of functional brain connectivity in moderating antidepressant treatment response in major depression. *American Journal of Psychiatry*, 177(2), 143–154. <https://doi.org/10.1176/appi.ajp.2019.18070870>, PubMed: 31537090
- Cook, I. A., Hunter, A. M., Caudill, M. M., Abrams, M. J., & Leuchter, A. F. (2020). Prospective testing of a neurophysiologic biomarker for treatment decisions in major depressive disorder: The PRISE-MD trial. *Journal of Psychiatric Research*, 124, 159–165. <https://doi.org/10.1016/j.jpsychires.2020.02.028>, PubMed: 32169689
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Costafreda, S. G., Khanna, A., Mourao-Miranda, J., & Fu, C. H. (2009). Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. *Neuroreport*, 20(7), 637–641. <https://doi.org/10.1097/WNR.0b013e3283294159>, PubMed: 19339907
- Crane, N. A., Jenkins, L. M., Bhaumik, R., Dion, C., Gowins, J. R., Mickey, B. J., ... Langenecker, S. A. (2017). Multidimensional prediction of treatment response to antidepressants with cognitive control and functional MRI. *Brain*, 140(2), 472–486. <https://doi.org/10.1093/brain/aww326>, PubMed: 28122876
- Dichter, G. S., Gibbs, D., & Smoski, M. J. (2015). A systematic review of relations between resting-state functional-MRI and treatment response in major depressive disorder. *Journal of Affective Disorders*, 172, 8–17. <https://doi.org/10.1016/j.jad.2014.09.028>, PubMed: 25451389
- Dinga, R., Schmaal, L., Penninx, B. W., van Tol, M. J., Veltman, D. J., van Velzen, L., ... Marquand, A. F. (2019). Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. (2017). *NeuroImage: Clinical*, 22, 101796. <https://doi.org/10.1016/j.nicl.2019.101796>, PubMed: 30935858
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., ... Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, 23(1), 28–38. <https://doi.org/10.1038/nm.4246>, PubMed: 27918562
- Dunlop, B. W., & Mayberg, H. S. (2014). Neuroimaging-based biomarkers for treatment selection in major depressive disorder. *Dialogues in Clinical Neuroscience*, 16(4), 479–490. <https://doi.org/10.31887/DCNS.2014.16.4/bdunlop>, PubMed: 25733953
- Endicott, J., Nee, J., Harrison, W., & Blumenthal, R. (1993). Quality of life enjoyment and satisfaction questionnaire: A new measure. *Psychopharmacology Bulletin*, 29(2), 321–326. PubMed: 8290681
- Erguzel, T. T., Ozekes, S., Gultekin, S., Tarhan, N., Sayar, G. H., & Bayram, A. (2015). Neural network based response prediction of rTMS in major depressive disorder using QEEG concordance. *Psychiatry Investigation*, 12(1), 61–65. <https://doi.org/10.4306/pi.2015.12.1.61>, PubMed: 25670947
- Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, 68(2), 118–124. <https://doi.org/10.1016/j.biopsych.2010.01.027>, PubMed: 20303067
- Fan, S., Nemati, S., Akiki, T. J., Roscoe, J., Averill, C. L., Fouda, S., ... Abdallah, C. G. (2020). Pretreatment brain connectome fingerprint predicts treatment response in major depressive disorder. *Chronic Stress*, 4, 2470547020984726. <https://doi.org/10.1177/2470547020984726>, PubMed: 33458556
- Farmus, L., Arpin-Cribbie, C. A., & Cribbie, R. A. (2019). Continuous predictors of pretest-posttest change: Highlighting the impact of the regression artifact. *Frontiers in Applied Mathematics and Statistics*, 4, 64. <https://doi.org/10.3389/fams.2018.00064>
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1), 1–47. <https://doi.org/10.1093/cercor/1.1.1>, PubMed: 1822724
- Finn, E. S. (2021). Is it time to put rest to rest? *Trends in Cognitive Sciences*, 25(12), 1021–1032. <https://doi.org/10.1016/j.tics.2021.09.005>, PubMed: 34625348
- Fischer, A. S., Holt-Gosselin, B., Fleming, S. L., Hack, L. M., Ball, T. M., Schatzberg, A. F., & Williams, L. M. (2021). Intrinsic reward circuit connectivity profiles underlying symptom and quality of life outcomes following antidepressant medication: A report from the iSPOT-D trial. *Neuropsychopharmacology*, 46(4), 809–819. <https://doi.org/10.1038/s41386-020-00905-3>, PubMed: 33230268
- Fonseka, T. M., MacQueen, G. M., & Kennedy, S. H. (2018). Neuroimaging biomarkers as predictors of treatment outcome in major depressive disorder. *Journal of Affective Disorders*, 233, 21–35. <https://doi.org/10.1016/j.jad.2017.10.049>, PubMed: 29150145
- Fonzo, G. A., Etkin, A., Zhang, Y., Wu, W., Cooper, C., Chin-Fatt, C., ... Trivedi, M. H. (2019). Brain regulation of emotional conflict predicts antidepressant treatment response for depression. *Nature Human Behaviour*, 3(12), 1319–1331. <https://doi.org/10.1038/s41562-019-0732-1>, PubMed: 31548678
- Fried, E. (2017). Moving forward: How depression heterogeneity hinders progress in treatment and research. *Expert Review of*



- Neurotherapeutics*, 17(5), 423–425. <https://doi.org/10.1080/14737175.2017.1307737>, PubMed: 28293960
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197. <https://doi.org/10.1016/j.jad.2016.10.019>, PubMed: 27792962
- Fried, E. I., Nesse, R. M., Zivin, K., Guille, C., & Sen, S. (2014). Depression is more than the sum score of its parts: Individual DSM symptoms have different risk factors. *Psychological Medicine*, 44(10), 2067–2076. <https://doi.org/10.1017/S0033291713002900>, PubMed: 24289852
- Frisch, M. B., Clark, M. P., Rouse, S. V., Rudd, M. D., Paweleck, J. K., Greenstone, A., & Kopplin, D. A. (2005). Predictive and treatment validity of life satisfaction and the quality of life inventory. *Assessment*, 12(1), 66–78. <https://doi.org/10.1177/1073191104268006>, PubMed: 15695744
- Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, 1(1), 13–36. <https://doi.org/10.1089/brain.2011.0008>, PubMed: 22432952
- Friston, K. J., Preller, K. H., Mathys, C., Cagnan, H., Heinzle, J., Razi, A., & Zeidman, P. (2019). Dynamic causal modelling revisited. *NeuroImage*, 199, 730–744. <https://doi.org/10.1016/j.neuroimage.2017.02.045>, PubMed: 28219774
- Frässle, S., Marquand, A. F., Schmaal, L., Dinga, R., Veltman, D. J., van der Wee, N. J., ... Stephan, K. E. (2020). Predicting individual clinical trajectories of depression with generative embedding. *NeuroImage: Clinical*, 26, 102213. <https://doi.org/10.1016/j.nicl.2020.102213>, PubMed: 32197140
- Frässle, S., Yao, Y., Schöbi, D., Aponte, E. A., Heinzle, J., & Stephan, K. E. (2018). Generative models for clinical applications in computational psychiatry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 9(3), e1460. <https://doi.org/10.1002/wcs.1460>, PubMed: 29369526
- Fu, C. H., Mourao-Miranda, J., Costafreda, S. G., Khanna, A., Marquand, A. F., Williams, S. C., & Brammer, M. J. (2008). Pattern classification of sad facial processing: Toward the development of neurobiological markers in depression. *Biological Psychiatry*, 63(7), 656–662. <https://doi.org/10.1016/j.biopsych.2007.08.020>, PubMed: 17949689
- Gilbert, J. R., Symmonds, M., Hanna, M. G., Dolan, R. J., Friston, K. J., & Moran, R. J. (2016). Profiling neuronal ion channelopathies with non-invasive brain imaging and dynamic causal models: Case studies of single gene mutations. *NeuroImage*, 124, 43–53. <https://doi.org/10.1016/j.neuroimage.2015.08.057>, PubMed: 26342528
- Gillan, C. M., & Whelan, R. (2017). What big data can do for treatment in psychiatry. *Current Opinion in Behavioral Sciences*, 18, 34–42. <https://doi.org/10.1016/j.cobeha.2017.07.003>
- Gläscher, J., Daw, N., Dayan, P., & O’Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585–595. <https://doi.org/10.1016/j.neuron.2010.04.016>, PubMed: 20510862
- Godlewska, B. R., Browning, M., Norbury, R., Igoumenou, A., Cowen, P. J., & Harmer, C. J. (2018). Predicting treatment response in depression: The role of anterior cingulate cortex. *International Journal of Neuropsychopharmacology*, 21(11), 988–996. <https://doi.org/10.1093/ijnp/pyy069>, PubMed: 30124867
- Goldberg, D. (2011). The heterogeneity of “major depression.” *World Psychiatry*, 10(3), 226–228. <https://doi.org/10.1002/j.2051-5545.2011.tb00061.x>, PubMed: 21991283
- Goldstein-Piekarski, A. N., Korgaonkar, M. S., Green, E., Suppes, T., Schatzberg, A. F., Hastie, T., ... Williams, L. M. (2016). Human amygdala engagement moderated by early life stress exposure is a biobehavioral target for predicting recovery on antidepressants. *Proceedings of the National Academy of Sciences*, 113(42), 11955–11960. <https://doi.org/10.1073/pnas.1606671113>, PubMed: 27791054
- Goldstein-Piekarski, A. N., Staveland, B. R., Ball, T. M., Yesavage, J., Korgaonkar, M. S., & Williams, L. M. (2018). Intrinsic functional connectivity predicts remission on antidepressants: A randomized controlled trial to identify clinically applicable imaging biomarkers. *Translational Psychiatry*, 8(1), 1–11. <https://doi.org/10.1038/s41398-018-0100-3>, PubMed: 29507282
- Greer, T. L., Kurian, B. T., & Trivedi, M. H. (2010). Defining and measuring functional recovery from depression. *CNS Drugs*, 24(4), 267–284. <https://doi.org/10.2165/11530230-000000000-00000>, PubMed: 20297853
- Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., ... Schatzberg, A. F. (2007). Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, 62(5), 429–437. <https://doi.org/10.1016/j.biopsych.2006.09.020>, PubMed: 17210143
- Habert, J., Katzman, M. A., Oluboka, O. J., McIntyre, R., McIntosh, D., MacQueen, G., & Kennedy, S. (2016). Functional recovery in major depressive disorder. *Primary Care Companion for CNS Disorders*, 18(5). <https://doi.org/10.4088/PCC.15r01926>, PubMed: 27835721
- Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, 6(4), 278–296. <https://doi.org/10.1111/j.2044-8260.1967.tb00530.x>, PubMed: 6080235
- Han, K.-M., Ham, B.-J., & Kim, Y.-K. (2021). Development of neuroimaging-based biomarkers in major depression. *Advances in Experimental Medicine and Biology*, 1305, 85–99. [https://doi.org/10.1007/978-981-33-6044-0\\_6](https://doi.org/10.1007/978-981-33-6044-0_6), PubMed: 33834396
- Hasanzadeh, F., Mohebbi, M., & Rostami, R. (2019). Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal. *Journal of Affective Disorders*, 256, 132–142. <https://doi.org/10.1016/j.jad.2019.05.070>, PubMed: 31176185
- Hofmann, S. G., Curtiss, J., Carpenter, J. K., & Kind, S. (2017). Effect of treatments for depression on quality of life: A meta-analysis. *Cognitive Behaviour Therapy*, 46(4), 265–286. <https://doi.org/10.1080/16506073.2017.1304445>, PubMed: 28440699
- Hopman, H., Chan, S., Chu, W., Lu, H., Tse, C.-Y., Chau, S., ... Neggers, S. (2021). Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. *Journal of Affective Disorders*, 290, 261–271. <https://doi.org/10.1016/j.jad.2021.04.081>, PubMed: 34010751
- Hough, C. M., Bersani, F. S., Mellon, S. H., Morford, A. E., Lindqvist, D., Reus, V. I., ... Wolkowitz, O. M. (2021). Pre-treatment allostatic load and metabolic dysregulation predict SSRI response in major depressive disorder: A preliminary report. *Psychological*

- Medicine*, 51(12), 2117–2125. <https://doi.org/10.1017/S0033291720000896>, PubMed: 32438932
- Huys, Q. J., Daw, N. D., & Dayan, P. (2015). Depression: A decision-theoretic analysis. *Annual Review of Neuroscience*, 38, 1–23. <https://doi.org/10.1146/annurev-neuro-071714-033928>, PubMed: 25705929
- Huys, Q. J., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3), 404–413. <https://doi.org/10.1038/nn.4238>, PubMed: 26906507
- Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood and Anxiety Disorders*, 3(1), 1–16. <https://doi.org/10.1186/2045-5380-3-12>, PubMed: 23782813
- IsHak, W. W., Balayan, K., Bresee, C., Greenberg, J. M., Fakhry, H., Christensen, S., & Rapaport, M. H. (2013). A descriptive analysis of quality of life using patient-reported measures in major depressive disorder in a naturalistic outpatient setting. *Quality of Life Research*, 22(3), 585–596. <https://doi.org/10.1007/s11136-012-0187-6>, PubMed: 22544416
- IsHak, W. W., Greenberg, J. M., Balayan, K., Kapitanski, N., Jeffrey, J., Fathy, H., ... Rapaport, M. H. (2011). Quality of life: The ultimate outcome measure of interventions in major depressive disorder. *Harvard Review of Psychiatry*, 19(5), 229–239. <https://doi.org/10.3109/10673229.2011.614099>, PubMed: 21916825
- James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., ... Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017. *Lancet*, 392(10159), 1789–1858. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7), PubMed: 30496104
- Janssen, R. J., Mourão-Miranda, J., & Schnack, H. G. (2018). Making individual prognoses in psychiatry using neuroimaging and machine learning. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(9), 798–808. <https://doi.org/10.1016/j.bpsc.2018.04.004>, PubMed: 29789268
- Jaworska, N., de la Salle, S., Ibrahim, M.-H., Blier, P., & Knott, V. (2019). Leveraging machine learning approaches for predicting antidepressant treatment response using electroencephalography (EEG) and clinical data. *Frontiers in Psychiatry*, 9, 768. <https://doi.org/10.3389/fpsy.2018.00768>, PubMed: 30692945
- Ju, Y., Horien, C., Chen, W., Guo, W., Lu, X., Sun, J., ... Li, L. (2020). Connectome-based models can predict early symptom improvement in major depressive disorder. *Journal of Affective Disorders*, 273, 442–452. <https://doi.org/10.1016/j.jad.2020.04.028>, PubMed: 32560939
- Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder: A meta-analysis of resting-state functional connectivity. *JAMA Psychiatry*, 72(6), 603–611. <https://doi.org/10.1001/jamapsychiatry.2015.0071>, PubMed: 25785575
- Kang, S.-G., & Cho, S.-E. (2020). Neuroimaging biomarkers for predicting treatment response and recurrence of major depressive disorder. *International Journal of Molecular Sciences*, 21(6), 2148. <https://doi.org/10.3390/ijms21062148>, PubMed: 32245086
- Karim, H. T., Wang, M., Andreescu, C., Tudorascu, D., Butters, M. A., Karp, J. F., ... Aizenstein, H. J. (2018). Acute trajectories of neural activation predict remission to pharmacotherapy in late-life depression. *NeuroImage: Clinical*, 19, 831–839. <https://doi.org/10.1016/j.nicl.2018.06.006>, PubMed: 30013927
- Katahira, K., & Toyama, A. (2021). Revisiting the importance of model fitting for model-based fMRI: It does matter in computational psychiatry. *PLoS Computational Biology*, 17(2), e1008738. <https://doi.org/10.1371/journal.pcbi.1008738>, PubMed: 33561125
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1), 1–9. <https://doi.org/10.1186/s12916-019-1426-2>, PubMed: 31665002
- Kennedy, S. H., Downar, J., Evans, K. R., Feilolter, H., Lam, R. W., MacQueen, G. M., ... Soares, C. (2012). The Canadian biomarker integration network in depression (CAN-BIND): Advances in response prediction. *Current Pharmaceutical Design*, 18(36), 5976–5989. <https://doi.org/10.2174/138161212803523635>, PubMed: 22681173
- Keren, H., O’Callaghan, G., Vidal-Ribas, P., Buzzell, G. A., Brotman, M. A., Leibenluft, E., ... Stringaris, A. (2018). Reward processing in depression: A conceptual and meta-analytic review across fMRI and EEG studies. *American Journal of Psychiatry*, 175(11), 1111–1120. <https://doi.org/10.1176/appi.ajp.2018.17101124>, PubMed: 29921146
- Khodayari-Rostamabad, A., Reilly, J. P., Hasey, G., de Bruin, H., & MacCrimmon, D. (2010). Using pre-treatment EEG data to predict response to SSRI treatment for MDD. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6103–6106). <https://doi.org/10.1109/IEMBS.2010.5627823>, PubMed: 21097134
- Khodayari-Rostamabad, A., Reilly, J. P., Hasey, G. M., de Bruin, H., & MacCrimmon, D. (2011). Using pre-treatment electroencephalography data to predict response to transcranial magnetic stimulation therapy for major depression. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6418–6421). <https://doi.org/10.1109/IEMBS.2011.6091584>, PubMed: 22255807
- Khodayari-Rostamabad, A., Reilly, J. P., Hasey, G. M., de Bruin, H., & MacCrimmon, D. J. (2013). A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder. *Clinical Neurophysiology*, 124(10), 1975–1985. <https://doi.org/10.1016/j.clinph.2013.04.010>, PubMed: 23684127
- Klöbl, M., Gryglewski, G., Rischka, L., Godbersen, G. M., Unterholzner, J., Reed, M. B., ... Lanzenberger, R. (2020). Predicting antidepressant citalopram treatment response via changes in brain functional connectivity after acute intravenous challenge. *Frontiers in Computational Neuroscience*, 14, 554186. <https://doi.org/10.3389/fncom.2020.554186>, PubMed: 33123000
- Kong, Y., Gao, S., Yue, Y., Hou, Z., Shu, H., Xie, C., ... Yuan, Y. (2021). Spatio-temporal graph convolutional network for diagnosis and treatment response prediction of major depressive disorder from functional connectivity. *Human Brain Mapping*, 42(12), 3922–3933. <https://doi.org/10.1002/hbm.25529>, PubMed: 33969930
- Korgaonkar, M. S., Goldstein-Piekarski, A. N., Fornito, A., & Williams, L. M. (2020). Intrinsic connectomes are a predictive biomarker of remission in major depressive disorder. *Molecular*

- Psychiatry*, 25(7), 1537–1549. <https://doi.org/10.1038/s41380-019-0574-2>, PubMed: 31695168
- Koshiyama, D., Kirihara, K., Usui, K., Tada, M., Fujioka, M., Morita, S., ... Kasai, K. (2020). Resting-state EEG beta band power predicts quality of life outcomes in patients with depressive disorders: A longitudinal investigation. *Journal of Affective Disorders*, 265, 416–422. <https://doi.org/10.1016/j.jad.2020.01.030>, PubMed: 32090768
- Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted cognitive processes in major depression: A predictive processing perspective. *Biological Psychiatry*, 87(5), 388–398. <https://doi.org/10.1016/j.biopsych.2019.07.017>, PubMed: 31515055
- Kupfer, D. J., Frank, E., & Phillips, M. L. (2012). Major depressive disorder: New clinical, neurobiological, and treatment perspectives. *Lancet*, 379(9820), 1045–1055. [https://doi.org/10.1016/S0140-6736\(11\)60602-8](https://doi.org/10.1016/S0140-6736(11)60602-8), PubMed: 22189047
- Lam, R. W., Parikh, S. V., Michalak, E. E., Dewa, C. S., & Kennedy, S. H. (2015). Canadian Network for Mood and Anxiety Treatments (CANMAT) consensus recommendations for functional outcomes in major depressive disorder. *Annals of Clinical Psychiatry*, 27(2), 142–149. PubMed: 25954941
- Leaver, A. M., Wade, B., Vasavada, M., Helleman, G., Joshi, S. H., Espinoza, R., & Narr, K. L. (2018). Fronto-temporal connectivity predicts ECT outcome in major depression. *Frontiers in Psychiatry*, 9, 92. <https://doi.org/10.3389/fpsy.2018.00092>, PubMed: 29618992
- Leon, A., Solomon, D., Mueller, T., Turvey, C., Endicott, J., & Keller, M. (1999). The range of impaired functioning tool (life–rift): A brief measure of functional impairment. *Psychological Medicine*, 29(4), 869–878. <https://doi.org/10.1017/S0033291799008570>, PubMed: 10473314
- Leuchter, A. F., Cook, I. A., Gilmer, W. S., Marangell, L. B., Burgoyne, K. S., Howland, R. H., ... Greenwald, S. (2009). Effectiveness of a quantitative electroencephalographic biomarker for predicting differential response or remission with escitalopram and bupropion in major depressive disorder. *Psychiatry Research*, 169(2), 132–138. <https://doi.org/10.1016/j.psychres.2009.04.004>, PubMed: 19709754
- Li, B., Friston, K., Mody, M., Wang, H., Lu, H., & Hu, D. (2018). A brain network model for depression: From symptom understanding to disease intervention. *CNS Neuroscience and Therapeutics*, 24(11), 1004–1019. <https://doi.org/10.1111/cns.12998>, PubMed: 29931740
- Liu, G. F., Lu, K., Mogg, R., Mallick, M., & Mehrotra, D. V. (2009). Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials? *Statistics in Medicine*, 28(20), 2509–2530. <https://doi.org/10.1002/sim.3639>, PubMed: 19610129
- Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 68(5), 304–305. <https://doi.org/10.1037/h0025105>, PubMed: 6062585
- Luedtke, A., & Kessler, R. C. (2021). New directions in research on heterogeneity of treatment effects for major depression. *JAMA Psychiatry*, 78(5), 478–480. <https://doi.org/10.1001/jamapsychiatry.2020.4489>, PubMed: 33595616
- Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A., Lamy, C., Magrou, L., Vezoli, J., ... Kennedy, H. (2014). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral Cortex*, 24(1), 17–36. <https://doi.org/10.1093/cercor/bhs270>, PubMed: 23010748
- Marquand, A. F., Mourão-Miranda, J., Brammer, M. J., Cleare, A. J., & Fu, C. H. (2008). Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport*, 19(15), 1507–1511. <https://doi.org/10.1097/WNR.0b013e328310425e>, PubMed: 18797307
- Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, 113(1), 181–190. <https://doi.org/10.1037/0033-2909.113.1.181>
- Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., ... Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, 45(5), 651–660. <https://doi.org/10.1016/j.neuron.2005.02.014>, PubMed: 15748841
- Meyer, B. M., Rabl, U., Huemer, J., Bartova, L., Kalcher, K., Provenzano, J., ... Pezawas, L. (2019). Prefrontal networks dynamically related to recovery from major depressive disorder: A longitudinal pharmacological fMRI study. *Translational Psychiatry*, 9(1), 64. <https://doi.org/10.1038/s41398-019-0395-8>, PubMed: 30718459
- Miller, J. M., Schneck, N., Siegle, G. J., Chen, Y., Ogden, R. T., Kikuchi, T., ... Parsey, R. V. (2013). fMRI response to negative words and SSRI treatment outcome in major depressive disorder: A preliminary study. *Psychiatry Research: Neuroimaging*, 214(3), 296–305. <https://doi.org/10.1016/j.pscychres.2013.08.001>, PubMed: 24446548
- Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134(4), 382–389. <https://doi.org/10.1192/bjp.134.4.382>, PubMed: 444788
- Moran, R. J., Symmonds, M., Stephan, K. E., Friston, K. J., & Dolan, R. J. (2011). An in vivo assay of synaptic function mediating human cognition. *Current Biology*, 21(15), 1320–1325. <https://doi.org/10.1016/j.cub.2011.06.053>, PubMed: 21802302
- Moreno-Ortega, M., Prudic, J., Rowny, S., Patel, G., Kangarlou, A., Lee, S., ... Javitt, D. C. (2019). Resting state functional connectivity predictors of treatment response to electroconvulsive therapy in depression. *Scientific Reports*, 9(1), 5071. <https://doi.org/10.1038/s41598-019-41175-4>, PubMed: 30911075
- Morton, E., Bhat, V., Giacobbe, P., Lou, W., Michalak, E. E., Chakrabarty, T., ... CAN-BIND Investigator Team. (2021). Impacts on quality of life with escitalopram monotherapy and aripiprazole augmentation in patients with major depressive disorder: A CAN-BIND report. *Pharmacopsychiatry*, 54(5), 225–231. <https://doi.org/10.1055/a-1385-0263>, PubMed: 33652477
- Mukherjee, D., Lee, S., Kazinka, R., Satterthwaite, T. D., & Kable, J. W. (2020). Multiple facets of value-based decision making in major depressive disorder. *Scientific Reports*, 10(1), 3415. <https://doi.org/10.1038/s41598-020-60230-z>, PubMed: 32099062
- Mumtaz, W., Xia, L., Mohd Yasin, M. A., Azhar Ali, S. S., & Malik, A. S. (2017). A wavelet-based technique to predict treatment outcome for major depressive disorder. *PLoS ONE*, 12(2), e0171409. <https://doi.org/10.1371/journal.pone.0171409>, PubMed: 28152063
- Must, A., Horvath, S., Nemeth, V. L., & Janka, Z. (2013). The Iowa gambling task in depression—What have we learned about

- sub-optimal decision-making strategies? *Frontiers in Psychology*, 4, 732. <https://doi.org/10.3389/fpsyg.2013.00732>, PubMed: 24133474
- Nemati, S., Akiki, T. J., Roscoe, J., Ju, Y., Averill, C. L., Fouda, S., ... Abdallah, C. G. (2020). A unique brain connectome fingerprint predates and predicts response to antidepressants. *iScience*, 23(1), 100800. <https://doi.org/10.1016/j.isci.2019.100800>, PubMed: 31918047
- Nguyen, K. P., Fatt, C. C., Treacher, A., Mellema, C., Trivedi, M. H., & Montillo, A. (2019). Predicting response to the antidepressant bupropion using pretreatment fMRI. In *International Workshop on Predictive Intelligence in Medicine* (pp. 53–62). [https://doi.org/10.1007/978-3-030-32281-6\\_6](https://doi.org/10.1007/978-3-030-32281-6_6), PubMed: 31709423
- O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104(1), 35–53. <https://doi.org/10.1196/annals.1390.022>, PubMed: 17416921
- Olbrich, S., & Arns, M. (2013). EEG biomarkers in major depressive disorder: Discriminative power and prediction of treatment response. *International Review of Psychiatry*, 25(5), 604–618. <https://doi.org/10.3109/09540261.2013.816269>, PubMed: 24151805
- Oluboka, O. J., Katzman, M. A., Habert, J., McIntosh, D., MacQueen, G. M., Milev, R. V., ... Blier, P. (2018). Functional recovery in major depressive disorder: Providing early optimal treatment for the individual patient. *International Journal of Neuropsychopharmacology*, 21(2), 128–144. <https://doi.org/10.1093/ijnp/pyx081>, PubMed: 29024974
- Patel, M. J., Andreescu, C., Price, J. C., Edelman, K. L., Reynolds, C. F., III, & Aizenstein, H. J. (2015). Machine learning approaches for integrating clinical and imaging features in LLD classification and response prediction. *International Journal of Geriatric Psychiatry*, 30(10), 1056–1067. <https://doi.org/10.1002/gps.4262>, PubMed: 25689482
- Paulus, M. P., & Angela, J. Y. (2012). Emotion and decision-making: Affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, 16(9), 476–483. <https://doi.org/10.1016/j.tics.2012.07.009>, PubMed: 22898207
- Pei, C., Sun, Y., Zhu, J., Wang, X., Zhang, Y., Zhang, S., ... Lu, Q. (2020). Ensemble learning for early-response prediction of antidepressant treatment in major depressive disorder. *Journal of Magnetic Resonance Imaging*, 52(1), 161–171. <https://doi.org/10.1002/jmri.27029>, PubMed: 31859419
- Perlman, K., Benrimoh, D., Israel, S., Rollins, C., Brown, E., Tunteng, J.-F., ... Berlim, M. T. (2019). A systematic meta-review of predictors of antidepressant treatment outcome in major depressive disorder. *Journal of Affective Disorders*, 243, 503–515. <https://doi.org/10.1016/j.jad.2018.09.067>, PubMed: 30286415
- Phillips, M. L., Chase, H. W., Sheline, Y. I., Etkin, A., Almeida, J. R., Deckersbach, T., & Trivedi, M. H. (2015). Identifying predictors, moderators, and mediators of antidepressant response in major depressive disorder: Neuroimaging approaches. *American Journal of Psychiatry*, 172(2), 124–138. <https://doi.org/10.1176/appi.ajp.2014.14010076>, PubMed: 25640931
- Pizzagalli, D. A. (2014). Depression, stress, and anhedonia: Toward a synthesis and integrated model. *Annual Review of Clinical Psychology*, 10, 393–423. <https://doi.org/10.1146/annurev-clinpsy-050212-185606>, PubMed: 24471371
- Queirazza, F., Fouragnan, E., Steele, J. D., Cavanagh, J., & Philiastides, M. G. (2019). Neural correlates of weighted reward prediction error during reinforcement learning classify response to cognitive behavioral therapy in depression. *Science Advances*, 5(7), eaav4962. <https://doi.org/10.1126/sciadv.aav4962>, PubMed: 31392266
- Rabinoff, M., Kitchen, C., Cook, I., & Leuchter, A. (2011). Evaluation of quantitative EEG by classification and regression trees to characterize responders to antidepressant and placebo treatment. *Open Medical Informatics Journal*, 5, 1–8. <https://doi.org/10.2174/1874431101105010001>, PubMed: 21603560
- Rajpurkar, P., Yang, J., Dass, N., Vale, V., Keller, A. S., Irvin, J., ... Williams, L. M. (2020). Evaluation of a machine learning model based on pretreatment symptoms and electroencephalographic features to predict outcomes of antidepressant treatment in adults with depression: A prespecified secondary analysis of a randomized clinical trial. *JAMA Network Open*, 3(6), e206653. <https://doi.org/10.1001/jamanetworkopen.2020.6653>, PubMed: 32568399
- Read, J., & Williams, J. (2018). Adverse effects of antidepressants reported by a large international cohort: Emotional blunting, suicidality, and withdrawal effects. *Current Drug Safety*, 13(3), 176–186. <https://doi.org/10.2174/1574886313666180605095130>, PubMed: 29866014
- Robinson, O. J., & Chase, H. W. (2017). Learning and choice in mood disorders: Searching for the computational parameters of anhedonia. *Computational Psychiatry*, 1(1), 208–233. [https://doi.org/10.1162/CPSPY\\_a\\_00009](https://doi.org/10.1162/CPSPY_a_00009), PubMed: 29400358
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- Rupprechter, S., Stankevicius, A., Huys, Q. J., Steele, J. D., & Seriès, P. (2018). Major depression impairs the use of reward values for decision-making. *Scientific Reports*, 8(1), 13798. <https://doi.org/10.1038/s41598-018-31730-w>, PubMed: 30218084
- Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., ... Keller, M. B. (2003). The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, 54(5), 573–583. [https://doi.org/10.1016/S0006-3223\(02\)01866-8](https://doi.org/10.1016/S0006-3223(02)01866-8), PubMed: 12946886
- Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., ... Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR\*D report. *American Journal of Psychiatry*, 163(11), 1905–1917. <https://doi.org/10.1176/appi.2006.163.11.1905>, PubMed: 17074942
- Schöbi, D., Homberg, F., Frässle, S., Endepols, H., Moran, R. J., Friston, K. J., ... Stephan, K. E. (2021). Model-based prediction of muscarinic receptor function from auditory mismatch negativity responses. *NeuroImage*, 237, 118096. <https://doi.org/10.1016/j.neuroimage.2021.118096>, PubMed: 33940149
- Schwab, J., Bialow, M., Clemmons, R., Martin, P., & Holzer, C. (1967). The Beck depression inventory with medical inpatients. *Acta Psychiatrica Scandinavica*, 43(3), 225–266. <https://doi.org/10.1111/j.1600-0447.1967.tb05762.x>, PubMed: 5585152

- Shahabi, M. S., Shalhaf, A., & Maghsoudi, A. (2021). Prediction of drug response in major depressive disorder using ensemble of transfer learning with convolutional neural network based on EEG. *Biocybernetics and Biomedical Engineering*, 41(3), 946–959. <https://doi.org/10.1016/j.bbe.2021.06.006>
- Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., ... Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences*, 106(6), 1942–1947. <https://doi.org/10.1073/pnas.0812686106>, PubMed: 19171889
- Sikora, M., Heffernan, J., Avery, E. T., Mickey, B. J., Zubietta, J.-K., & Peciña, M. (2016). Salience network functional connectivity predicts placebo effects in major depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1(1), 68–76. <https://doi.org/10.1016/j.bpsc.2015.10.002>, PubMed: 26709390
- Steffen, A., Nübel, J., Jacobi, F., Bätzing, J., & Holstiege, J. (2020). Mental and somatic comorbidity of depression: A comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. *BMC Psychiatry*, 20(1), 1–15. <https://doi.org/10.1186/s12888-020-02546-8>, PubMed: 32228541
- Stephan, K. E., & Friston, K. J. (2010). Analyzing effective connectivity with functional magnetic resonance imaging. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(3), 446–459. <https://doi.org/10.1002/wcs.58>, PubMed: 21209846
- Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A., Paliwal, S., Gard, T., ... Petzschner, F. H. (2016). Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience*, 10, 550. <https://doi.org/10.3389/fnhum.2016.00550>, PubMed: 27895566
- Stephan, K. E., Schlagenhaut, F., Huys, Q. J., Raman, S., Aponte, E. A., Brodersen, K. H., ... Heinz, A. (2017). Computational neuroimaging strategies for single patient predictions. *NeuroImage*, 145, 180–199. <https://doi.org/10.1016/j.neuroimage.2016.06.038>, PubMed: 27346545
- Stuhrmann, A., Suslow, T., & Dannlowski, U. (2011). Facial emotion processing in major depression: A systematic review of neuroimaging findings. *Biology of Mood and Anxiety Disorders*, 1(1), 1–17. <https://doi.org/10.1186/2045-5380-1-10>, PubMed: 22738433
- Sun, H., Jiang, R., Qi, S., Narr, K. L., Wade, B. S., Upston, J., ... Sui, J. (2020). Preliminary prediction of individual response to electroconvulsive therapy using whole-brain functional magnetic resonance imaging data. *NeuroImage: Clinical*, 26, 102080. <https://doi.org/10.1016/j.nicl.2019.102080>, PubMed: 31735637
- Symmonds, M., Moran, C. H., Leite, M. I., Buckley, C., Irani, S. R., Stephan, K. E., ... Moran, R. J. (2018). Ion channels in EEG: Isolating channel dysfunction in NMDA receptor antibody encephalitis. *Brain*, 141(6), 1691–1702. <https://doi.org/10.1093/brain/awy107>, PubMed: 29718139
- Taylor, J. J., Kurt, H. G., & Anand, A. (2021). Resting state functional connectivity biomarkers of treatment response in mood disorders: A review. *Frontiers in Psychiatry*, 12, 565136. <https://doi.org/10.3389/fpsy.2021.565136>, PubMed: 33841196
- Tian, S., Sun, Y., Shao, J., Zhang, S., Mo, Z., Liu, X., ... Lu, Q. (2020). Predicting escitalopram monotherapy response in depression: The role of anterior cingulate cortex. *Human Brain Mapping*, 41(5), 1249–1260. <https://doi.org/10.1002/hbm.24872>, PubMed: 31758634
- Tozzi, L., Goldstein-Piekarski, A. N., Korgaonkar, M. S., & Williams, L. M. (2020). Connectivity of the cognitive control network during response inhibition as a predictive and response biomarker in major depression: Evidence from a randomized clinical trial. *Biological Psychiatry*, 87(5), 462–472. <https://doi.org/10.1016/j.biopsych.2019.08.005>, PubMed: 31601424
- Trivedi, M. H., McGrath, P. J., Fava, M., Parsey, R. V., Kurian, B. T., Phillips, M. L., ... Weissman, M. M. (2016). Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. *Journal of Psychiatric Research*, 78, 11–23. <https://doi.org/10.1016/j.jpsychires.2016.03.001>, PubMed: 27038550
- Tu, Y.-K., Gunnell, D., & Gilthorpe, M. S. (2008). Simpson's paradox, Lord's paradox, and suppression effects are the same phenomenon—The reversal paradox. *Emerging Themes in Epidemiology*, 5(1), 1–9. <https://doi.org/10.1186/1742-7622-5-2>, PubMed: 18211676
- Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLoS ONE*, 14(11), e0224365. <https://doi.org/10.1371/journal.pone.0224365>, PubMed: 31697686
- Vai, B., Bulgarelli, C., Godlewska, B. R., Cowen, P. J., Benedetti, F., & Harmer, C. J. (2016). Fronto-limbic effective connectivity as possible predictor of antidepressant response to SSRI administration. *European Neuropsychopharmacology*, 26(12), 2000–2010. <https://doi.org/10.1016/j.euroneuro.2016.09.640>, PubMed: 27756525
- Vandeleur, C. L., Fassassi, S., Castela, E., Glaus, J., Strippoli, M.-P. F., Lasserre, A. M., ... Preisig, M. (2017). Prevalence and correlates of DSM-5 major depressive and related disorders in the community. *Psychiatry Research*, 250, 50–58. <https://doi.org/10.1016/j.psychres.2017.01.060>, PubMed: 28142066
- van den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology*, 20(8), 519–534. <https://doi.org/10.1016/j.euroneuro.2010.03.008>, PubMed: 20471808
- van der Vinne, N., Vollebregt, M. A., Rush, A. J., Eebes, M., van Putten, M. J., & Arns, M. (2021). EEG biomarker informed prescription of antidepressants in MDD: A feasibility trial. *European Neuropsychopharmacology*, 44, 14–22. <https://doi.org/10.1016/j.euroneuro.2020.12.005>, PubMed: 33509659
- Van Waarde, J., Scholte, H., Van Oudheusden, L., Verwey, B., Denys, D., & Van Wingen, G. (2015). A functional MRI marker may predict the outcome of electroconvulsive therapy in severe and treatment-resistant depression. *Molecular Psychiatry*, 20(5), 609–614. <https://doi.org/10.1038/mp.2014.78>, PubMed: 25092248
- Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7(1), 1–8. <https://doi.org/10.1186/1471-2105-7-91>, PubMed: 16504092
- Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*, 180, 68–77. <https://doi.org/10.1016/j.neuroimage.2017.06.061>, PubMed: 28655633
- 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*, 145, 166–179. <https://doi.org/10.1016/j.neuroimage.2016.10.038>, PubMed: 27989847

- Williams, L. M., Korgaonkar, M. S., Song, Y. C., Paton, R., Eagles, S., Goldstein-Piekarski, A., ... Etkin, A. (2015). Amygdala reactivity to emotional faces in the prediction of general and medication-specific responses to antidepressant treatment in the randomized iSPOT-D trial. *Neuropsychopharmacology*, *40*(10), 2398–2408. <https://doi.org/10.1038/npp.2015.89>, PubMed: 25824424
- Williams, L. M., Rush, A. J., Koslow, S. H., Wisniewski, S. R., Cooper, N. J., Nemeroff, C. B., ... Gordon, E. (2011). International study to predict optimized treatment for depression (iSPOT-D), a randomized clinical trial: Rationale and protocol. *Trials*, *12*(1), 1–17. <https://doi.org/10.1186/1745-6215-12-4>, PubMed: 21208417
- Wright, R. E. (1995). Logistic regression. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 217–244). American Psychological Association.
- Wu, C.-T., Dillon, D. G., Hsu, H.-C., Huang, S., Barrick, E., & Liu, Y.-H. (2018). Depression detection using relative EEG power induced by emotionally positive images and a conformal kernel support vector machine. *Applied Science*, *8*(8), 1244. <https://doi.org/10.3390/app8081244>
- Wu, W., Zhang, Y., Jiang, J., Lucas, M. V., Fonzo, G. A., Rolle, C. E., ... Etkin, A. (2020). An electroencephalographic signature predicts antidepressant response in major depression. *Nature Biotechnology*, *38*(4), 439–447. <https://doi.org/10.1038/s41587-019-0397-3>, PubMed: 32042166
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. <https://doi.org/10.1177/1745691617693393>, PubMed: 28841086
- Zhdanov, A., Atluri, S., Wong, W., Vaghei, Y., Daskalakis, Z. J., Blumberger, D. M., ... Farzan, F. (2020). Use of machine learning for predicting escitalopram treatment outcome from electroencephalography recordings in adult patients with depression. *JAMA Network Open*, *3*(1), e1918377. <https://doi.org/10.1001/jamanetworkopen.2019.18377>, PubMed: 31899530
- Zimmerman, M., McGlinchey, J. B., Posternak, M. A., Friedman, M., Attiullah, N., & Boerescu, D. (2006). How should remission from depression be defined? The depressed patient's perspective. *American Journal of Psychiatry*, *163*(1), 148–150. <https://doi.org/10.1176/appi.ajp.163.1.148>, PubMed: 16390903