doi:10.1111/pcn.12502

Psychiatry and Clinical Neurosciences

PCN Frontier Review

Computational neuroscience approach to biomarkers and treatments for mental disorders

Noriaki Yahata, PhD,[1,2,3] Kiyoto Kasai, MD, PhD[4] and Mitsuo Kawato, PhD[3]*

[1]*Department of Youth Mental Health, Graduate School of Medicine, The University of Tokyo, Tokyo,* [2]*Department of Molecular Imaging and Theranostics, National Institute of Radiological Sciences, National Institutes for Quantum and Radiological Science and Technology, Chiba,* [3]*ATR Brain Information Communication Research Laboratory Group, Kyoto,* [4]*Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan*

Psychiatry research has long experienced a stagnation stemming from a lack of understanding of the neurobiological underpinnings of phenomenologically defined mental disorders. Recently, the application of computational neuroscience to psychiatry research has shown great promise in establishing a link between phenomenological and pathophysiological aspects of mental disorders, thereby recasting current nosology in more biologically meaningful dimensions. In this review, we highlight recent investigations into computational neuroscience that have undertaken either theory- or data-driven approaches to quantitatively delineate the mechanisms of mental disorders. The theory-driven approach, including reinforcement learning models, plays an integrative role in this process by enabling correspondence between behavior and disorder-specific alterations at multiple levels of brain organization, ranging from molecules to cells to circuits. Previous studies have explicated a plethora of defining symptoms of mental disorders, including anhedonia, inattention, and poor executive function. The data-driven approach, on the other hand, is an emerging field in computational neuroscience seeking to identify disorder-specific features among high-dimensional big data. Remarkably, various machine-learning techniques have been applied to neuroimaging data, and the extracted disorder-specific features have been used for automatic case–control classification. For many disorders, the reported accuracies have reached 90% or more. However, we note that rigorous tests on independent cohorts are critically required to translate this research into clinical applications. Finally, we discuss the utility of the disorder-specific features found by the data-driven approach to psychiatric therapies, including neurofeedback. Such developments will allow simultaneous diagnosis and treatment of mental disorders using neuroimaging, thereby establishing 'theranostics' for the first time in clinical psychiatry.

Key words: biomarkers, computational psychiatry, machine learning, neuroimaging, resting-state functional connectivity.

Correspondence: Mitsuo Kawato, PhD, Computational Neuroscience Laboratories, Advanced Telecommunications Research Institute International, 2-2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan. Email: kawato@atr.jp
Received 31 August 2016; revised 19 December 2016; accepted 25 December 2016.

DESPITE LONG-STANDING AND ever-expanding efforts in clinical and basic research, patients with mental disorders have not fully benefitted from their outcome.[1,2] This unfortunate situation is reflected by the fact that mental disorders have ranked high in public health statistics, such as

disability-adjusted life years (DALY)[3] and global economic burden of illness.[4] There are a lack of effective biomarkers that facilitate the early detection and unambiguous diagnosis of disorders, and an absence of effective pharmacological agents that lead to efficient recovery or prevent the relapse of disorders. A common criticism of psychiatric health care relates to the current symptom-based, operationalized definition of mental disorders, as standardized in the DSM[5] and ICD.[6] Here, each mental disorder is assumed to be distinct and defined by a constellation of non-specific symptoms or syndromes, with no concrete biological underpinnings. In practice, however, clinicians face many exceptions to this phenomenological stratification – similar symptoms are often present across multiple disorders, and patients commonly have psychiatric comorbidities. The reification of a given mental disorder is highly heterogeneous,[2,7] suggesting that a single diagnosis may entail multiple pathogenic and pathophysiological mechanisms. Thus, biological research aiming to elucidate the etiology and pathophysiology of mental disorders has been obscured by the 'lens' of symptom-based stratification,[2] exemplified by the DSM and ICD, impeding substantial progress over the course of decades of psychiatry research.

To overcome this stagnation, growing attention has now been granted to a dimensional view aiming to recast mental disorders within a framework of biologically grounded, quantifiable constructs.[8] In this fiducial space, varying contributions of constructs (i.e., risk factors for disorders) constitute the spectrum of mental disorders within which current diagnoses, such as schizophrenia and mood disorder, are represented as entities with no definitive mutual boundaries. In this regard, the Research Domain Criteria (RDoC) project,[9] launched in 2009 by the US National Institute of Mental Health, has taken a leading role in identifying the constructs or 'biosignatures'[10] that encompass both normal and abnormal states of the brain circuitry. At present, the RDoC incorporates five domains of brain function (negative valence, positive valence, cognitive, social processes, and arousal/regulatory systems) to be characterized at multiple levels of brain organization, from genes, molecules, cells, and circuits to physiology in conjunction with an individual's behavior and self-report. On the other hand, apparently beyond the scope of the RDoC is how findings at each level can be aggregated to optimally elicit a comprehensive picture of the mechanisms of a mental disorder.[11] Thus, a complementary framework is required to disentangle the complex interactions among multiple levels and physical and social environments,[12] thereby achieving optimal integration of the multilevel findings.

In this review, we focus on the recent application of computational neuroscience in psychiatry research and discuss its potential role in: (i) revealing the biological underpinnings of mental disorders; (ii) deriving a novel framework for making objective, quantitative diagnoses; and (iii) developing treatment regimens in a clinical setting. The principal goal of computational neuroscience is to establish a mechanistic model of the brain that provides an integrative account for human behavior. Such a model can take either a theory- or data-driven approach.[12] Theory-driven approaches, including popular reinforcement learning (RL)[13] and game-theoretic models,[14] have been applied in psychiatry research[15,16] and provide a quantitative description of the mechanisms underlying aberrations of emotion, decision-making, executive function, and other functions. Data-driven approaches, on the other hand, have emerged more recently in psychiatry research, invoking growing attention, especially in the context of potential clinical applications.[12,17,18] The methodological framework of data-driven approaches typically involves a machine-learning technique, which is by nature disease non-specific and thus more amenable to general clinical usage. To elaborate on data-driven approaches, we highlight recent functional connectivity investigations,[17] demonstrating the increasing interest in characterizing mental disorders as disruptions of brain connections, and the recent availability of multi-center neuroimaging data that allow the application of various machine-learning techniques to extract disease-specific brain features for objective diagnosis and treatment.

THEORY-BASED APPROACH AS A PROBE INTO MECHANISMS UNDERLYING SYMPTOMS OF MENTAL DISORDERS

The goal of computational neuroscience is to establish an integrative, mechanistic model of our cognitive, emotional, and social processes.[19] Such a model would provide insights into the causal relationships between multiple levels of the brain's organization; for example, how the deficit or excess of a particular neurotransmitter drives to a particular

behavior. The emerging field of computational psychiatry can be viewed as an extension of computational neuroscience in terms of how models derived from a healthy, typically developed population can be extrapolated, or must be modified, to explain aberrant behaviors observed among patients with mental disorders.

Connectionist models, or parallel distributed processing (PDP) models, are the most classical approach in computational modeling that aim to conceptualize human perceptual and cognitive functions via an artificial neural network (i.e., the weighted interconnection of simple neuron-like processing units).[20] The information within the network is represented either in a single unit or by a pattern of activation over multiple units. Inputs into the network are processed through the propagation of activations in the units, and in many connectionist network models, the whole system eventually settles in the local minima of its energy state (attractor states).[21] The connection weights then trace the history of activation, representing the long-term memory of learning through experience.[21] Since its inception in the mid-1980s, the PDP framework has been successfully applied for modeling the acquisition of regular and irregular English tenses,[22] speech production,[23] and other brain functions, such as cognitive control[24] and consciousness.[25] In psychiatry research, attempts have been made to interpret various symptoms of schizophrenia using the PDP framework.[26] Previous studies have found abnormally excessive pruning of cortical connections in schizophrenia, and the computational model of this phenomenon has identified the emergence of a network pathology called a 'parasitic focus,' wherein the portion of neural modules was locked into a certain cognitive output irrespective of the state of the rest of the system.[26,27] This aberrant mechanism was thought to underlie symptoms of schizophrenia, such as thought insertion, thought broadcasting, and auditory/visual hallucinations.[26] Other characteristics associated with schizophrenia, such as impaired cognitive flexibility and control, were also considered in a connectionist model that explicitly incorporated the prefrontal cortex as a distinct module, which was demonstrated to be impaired.[28]

Decision-making has been a primary subject of investigation in computational psychiatry because aberrant decision-making constitutes the defining symptoms of many psychiatric conditions.[19] In this regard, RL models provide a powerful framework that relates the values of decision-making options to preceding actions and environments. More specifically, RL is concerned with the role of dopamine and its neural circuits in the basal ganglia and how reward-seeking and punishment-avoiding behaviors are affected by the presence (e.g., amount and timing) of reinforcers (reward and punishment). Such behaviors are directly connected to one's survival and constitute the principal concern of neural circuits.[15] A detailed description of RL, including model-based and model-free RL, is beyond the scope of this review. Instead, here we note that previous studies have associated aspects of RL with a wide variety of symptoms of mental disorders, including: (i) positive symptoms (e.g., delusions and hallucinations) in schizophrenia with abnormal reward prediction error,[29,30] aberrant activity in response to neural stimuli in the limbic motivational systems,[31] and abnormal levels of incentive salience[32] due to abnormally elevated levels of dopaminergic neurotransmission;[33] (ii) negative symptoms in schizophrenia with failure to represent the expected value of rewards[34–36] and abnormal effect–cost computations;[37] (iii) anhedonia in depression with a loss of reward sensitivity in a manner distinct from that involving dopaminergic transmission;[38] and (iv) impulsivity in attention-deficit hyperactivity disorder (ADHD) with impaired delay aversion caused by excessive discounting of delayed rewards.[39,40] For more detail, see other dedicated reviews.[41,42]

Social behavior is another important subject of investigation in computational psychiatry. Inferring the intentions and affective states of others is crucial for appropriate behavior and survival in social environments. Impairments in social functioning severe enough to impact quality of life are observed in many mental disorders, including schizophrenia, autism spectrum disorder (ASD), and personality disorders.[43] Game-theoretic approaches have been extensively used to investigate interpersonal decision-making that involves social fairness, strategic cooperation, and competition.[44] For example, in a stag-hunt game, two players (a human and a computerized agent) are required to cooperate to hunt for a high-payoff mobile stag (a male deer) instead of a low-payoff stationary rabbit.[45] A model-based behavioral analysis allows for an estimation of the recursion depth of inference about the co-player's intention. A previous study revealed that individuals with ASD with more severe symptoms were less

likely to exhibit a recursive inference.[46] Other game-theoretic investigations use a multiround trust game to probe normal and aberrant behaviors in reciprocal interactions. In this game, two players are endowed with a certain amount of money and play multiple rounds of monetary exchange according to the following rule: one player (an investor) entrusts some fraction of money in hand to the other player (a trustee) who actually receives triple the amount. The trustee then decides what fraction of the received money to transfer back to the investor. For optimal results, both players are required to cooperate by mentalizing the impact of one's actions on the other. Disorder-specific play styles have been observed in many mental disorders, including schizophrenia and other psychotic disorders,[47] major depressive disorder (MDD) and bipolar disorder,[48] ASD,[49] ADHD,[50] and borderline personality disorder.[51] Intriguingly, a Bayesian clustering approach yielded a pattern of clusters of healthy investors which, when investors were playing with trustees with mental disorders, significantly overlapped with the pattern correlating with the pathology of the trustees (i.e., ASD, ADHD, MDD, and borderline personality disorder).[52] These and other games derived from behavioral economics have been extensively employed in psychiatry research, providing a unique opportunity to phenotype individuals according to their behavior in social decision-making.

Lastly, we note that functional neuroimaging has been pivotal in deepening our understanding of the neural substrates of the modeled and phenotyped behaviors described in this section. Continuing efforts across multiple techniques will enable the multilevel integration of disorder-specific findings, contributing critically to establishing the objective and quantitative nosology of mental disorders in the future.

DATA-DRIVEN APPROACH FOR MAKING PREDICTIONS IN CLINICAL SETTINGS

For nearly two decades, neuroimaging has been an exploratory tool to identify structural, functional, and metabolic manifestations of mental disorders. Neuroimaging identifies features that may have a pivotal role in establishing a disorder's endophenotype, bridging its symptomatology and pathophysiology. A multitude of disorder-specific characteristics of the brain have been documented at the group level (i.e., case vs control), such as: regional gray matter volume[53–56] and cortical thickness[57–61] measured in high-resolution structural magnetic resonance (MR) images; white matter organization measured by diffusion tensor imaging;[62–66] key metabolites measured in MR spectroscopy;[67–69] positron emission tomography;[70–74] patterns of activation associated with psychological tasks and functional connectivity in participants at rest during functional magnetic resonance imaging (fMRI),[17,18] near-infrared spectroscopy,[75] and electroencephalogram.[76] While these disorder-specific alterations certainly contribute to the quantitative delineation of mental disorders, such findings have not led to the development of clinical methods to help determine the diagnostic and prognostic status of an individual. One major reason for this lack of development is that the effect size of the reported case–control separations is insufficient to permit classification on an individual basis.[18] Thus, a more sophisticated methodological framework is needed instead of the current simple thresholding to the observed case–control relationships.[18]

In the past several years, a surge of data-driven computational neuroscience studies has applied machine-learning techniques to neuroimaging data to uncover disorder-specific structural and functional features of the brain. Such hypothesis-free studies led to the formation of classifiers that objectively and automatically assign a diagnostic label (case or control) to an individual. To date, a variety of classification schemes have been proposed, and the accuracy of classification has reached as high as 90% for many disorders, including schizophrenia, MDD, and ASD; see a recent comprehensive survey of previous machine-learning studies that reported the construction of neuroimaging-based classifiers.[18,77,78] To lay people, it may appear as though the advent of artificial intelligence in the psychiatric clinic has arrived and will soon be responsible for diagnostic and therapeutic decision-making. In the present review, such naïve optimism is discouraged, and it is instead argued that several critical technical hurdles must be cleared before machine-learning-based classification can complement current psychiatric medicine.

Case–control classification based on the resting-state functional connectivity of the brain

In the following paragraphs, we focus on previous studies that have applied machine-learning

techniques to resting-state functional-connectivity magnetic resonance imaging (rs-fcMRI) data. The investigations are based on measuring spontaneous brain activity while subjects are at rest in an MR scanner without behavioral manipulations. Functional connectivity describes the similarity of low-frequency (<0.08–0.1 Hz) blood-oxygenation level dependent (BOLD) signals between voxels or groups of voxels (e.g., anatomical regions), typically measured by Pearson correlations. More recently, connectivity within and between functional networks, such as default mode and sensorimotor networks, has also become a target of investigation.[79,80] Functional connectivity can be measured either statically using the whole time series or dynamically within a predefined narrow time-window sliding across the time series. The pattern of low-frequency temporal correlations has been thought to reflect a personal history of Hebbian co-activation, as evidenced by previous reports on practice-induced changes in functional connectivity.[81] Currently, rs-fcMRI is an important tool for investigating the (dis)integrity of healthy and pathological brains of young and adult individuals,[17,82] as well as of model animals in preclinical studies. Thus, rs-fcMRI is a potential means of facilitating translational psychiatry research.

Various machine-learning techniques have been applied to mass rs-fcMRI data to objectively identify disorder-specific abnormalities in mental disorders, with which automatic case–control classifications were employed. Table 1 summarizes the previous attempts for schizophrenia,[80,83–98] MDD,[99–102] ADHD,[103–106] and ASD.[107–117] The take-home messages of this summary are as follows: (i) irrespective of disorder type, classification accuracy is, overall, 80–90%, comparable to those based on structural MRI data;[18,78] (ii) in many studies, especially for schizophrenia and MDD, the sample per group (case or control) is typically comprised of fewer than 100 participants; (iii) for all schizophrenia and MDD studies, the imaging data were acquired at a single site, whereas for many ADHD and ASD studies, the imaging data came from multiple sites, thanks to the recent multicenter imaging campaigns for these disorders;[115,116] (iv) inter-regional functional connectivity and the associated graph metrics are popular features used for classification; (v) head motions during scanning have been known to introduce artifacts in the functional connectivity estimate,[118,119] the effects of which are controlled by regression, masking (scrubbing),[118] or independent component analysis;[120] (vi) BOLD signal fluctuations of non-neuronal origins, such as respiration and cardiac activity, are removed by regressing out the signals in white matter and cerebrospinal fluid, although further inclusion of global signal fluctuation into the regressor is not unanimous among the studies due to the recent controversy;[121,122] (vii) support vector machine (SVM) and its variants are popular prediction methods, although some studies use classifiers with embedded regularization frameworks, such as least absolute shrinkage and selection operator (LASSO); (viii) - leave-one-out and k-fold cross-validation procedures are popular methods for model evaluation; and (ix) for all but one study,[114] the generalization capability of a classification scheme is untested in an independent cohort. We will elaborate more on this issue below.

Generalization of classification scheme

Among the properties above, a crucial factor that has hampered the clinical application of the machine-learning classification scheme is the lack of studies on their generalizability (i.e., whether or not a classification is effective in an independent population). The development of a classifier, including both the determination of a model form and the estimation of the associated parameters, proceeds with 'training' data, whereas the generalizability of a classifier must be evaluated using independent 'test' data. A rigorous assessment of a classifier's reliability on independent test data is particularly essential while determining clinical applicability. This stipulation is a harsh requirement for most previous neuroimaging studies because such studies typically rely on a small sample of participants that are recruited and scanned at a single imaging site. Although more recent multicenter neuroimaging projects have been ambitious in their target sample size, involving a few thousand or more participants, the issue of generalizability remains to be addressed.

Broadly speaking, generalization can be considered as two technical issues, namely overfitting and the treatment of nuisance variables (NV, to be described in the next section). Overfitting is a condition where a developed model perfectly describes the entire aspect of the training data, including the underlying relationship and associated noise, thereby making the fitting error asymptotically zero.[123,124] The noise characteristics are most likely

Table 1. Summary of previous developments in resting-state functional-connectivity-based classifiers for schizophrenia, major depressive disorder, attention-deficit hyperactivity disorder, and autism spectrum disorder

		Number of participants (male/female) Mean age ± SD (if available)									Classification performance			
Disorder	Reference	Patients	Controls	Other	Number of imaging sites	NV	Features	Dimensionality reduction & feature selection	Classifier	Number of features	Model evaluation	Accuracy (%)	Sensitivity (%)	Specificity (%)	Remarks
SCZ	Shen *et al.*[83]	32 (25/7) 24.0 ± 5.7	20 (16/4) 24.6 ± 4.5	—	1	HM	FC among 116 AAL ROI	Kendall's τ RC & LLE	C-means clustering	12	LOOCV	86.5	93.8	75.0	—
SCZ	Fan *et al.*[84]	31 (17/14) 24 ± 6	31 (16/15) 26 ± 4	—	1	—	FC patterns (ICA cmpnts)	ICA	SVM	6	Bagging	85.5	83.9	87.1	—
SCZ	Bassett *et al.*[85]	29 (18/11) 41.3 ± 9.3	29 (18/11) 41.1 ± 10.6	—	1	HM	Graph metric for FC among 90 AAL ROI	—	SVM	N/A	Random sampling	75	85	64	—
SCZ	Tang *et al.*[86]	22 (15/7) 24.5 ± 6.7	22 (15/7) 26.1 ± 6.5	—	1	HM, GS, WM & CSF	FC among 90 ROI from WFU PickAtlas	Kendall's τ RC & PCA	SVM	550	LOOCV	93.2	86.4	100.0	—
SCZ	Venkataraman *et al.*[87]	18 (18/0) 41.6 ± 9.5	18 (18/0) 39.2 ± 10.8	—	1	WM, VT & GS	FC among 77 ROI	Gini importance	RF	1096	(RF)	75	—	—	—
SCZ	Anderson and Cohen[88]	72 (58/14) 38.2 ± 13.9	74 (51/23) 35.8 ± 11.6	—	1	—	Graph–theoretic connectivity measures	ICA	SVM	13	10-fold CV	65	—	—	Data adopted from COBRE[†]
SCZ	Arbabshirani *et al.*[89]	28 39.7±10.1	28 36.5±11.3		1	—	Correlation among ICA cmpnts	PCA & MDL criterion for ICA	KNN	36	LOOCV	96	100	92	—
SCZ	Fekete *et al.*[90]	8 (8/0) 44.1±11.4	10 (10/0) 41.8±11.2	—	1	HM, HM squared, WM, CSF & GS	CN measures for FC among 116 AAL ROI	Thresholding (*t*-statistics) & BDopt	SVM	12	LOOCV	100	100	100	—
SCZ	Su *et al.*[91]	32 (25/7) 24.0 ± 5.7	32 (23/9) 25.0 ± 4.5	—	1	HM	Extended MIC from FC among 116 AAL ROI	—	SVM	6670	LOOCV	82.8	84.4	81.2	—
SCZ/MDD	Yu *et al.*[92]	32 (25/7) 24.0 ± 5.7	38 (27/11) 24.4 ± 4.5	MDD 19 (11/8) 26.7 ± 7.6	1	HM & GS	FC among 116 AAL ROI	IDA	SVM	∼330	LOOCV	80.9	81.3	78.9	Sensitivity and specificity are from the diagonal elements of the 3-group confusion matrix
SCZ	Yu *et al.*[93]	24 (12/12) 25.4 ± 6.3	23 (12/11) 25.5 ± 5.5	uSIB 25 (15/10) 25.6 ± 6.8	1	HM & GS	FC among 116 AAL ROI	PCA	SVM	330	LOOCV	62.0	66.7	63.6	Sensitivity and specificity are from the diagonal elements of the 3-group confusion matrix

Table 1. (*Continued*)

Disorder	Reference	Number of participants (male/female) Mean age ± SD (if available)			Number of imaging sites	NV	Features	Dimensionality reduction & feature selection	Classifier	Number of features	Model evaluation	Classification performance			Remarks
		Patients	Controls	Other								Accuracy (%)	Sensitivity (%)	Specificity (%)	
SCZ	Watanabe *et al.*[94]	54 (48/6) 35.5 ± 13.1	67 (46/21) 35.2 ± 11.7	—	1	HM, dHM, WM & CSF	FC among 347 nodes across the brain	(→)	Elastic net / GraphNet / Fused LASSO	3076/ 3403/ 3140	10-fold CV	73.5/ 70.3/ 71.9	—	—	Data adopted from COBRE†
SCZ	Cheng *et al.*[95]	19 (13/6) 33.1 ± 10.9	29 (14/15) 28.1 ± 8.4	—	1	HM, WM & CSF	Nodal BC of 278 FC ROI	Thresholding (BC)	SVM	10	LOOCV	79	74	83	—
SCZ	Chyzhyk *et al.*[96]	72 (58/24) 38.1	74 (51/23) 35.8	—	1	WM, CSF & PCA cmpnts	ReHo / ALFF	Pearson's correlation	V-ELM	2113/ 2198	10-fold CV	84.2/87.7	—	—	Data adopted from COBRE†
SCZ	Kaufmann *et al.*[97]	71 (42/29) 28.2 ± 7.8	196 (114/82) 31.5 ± 7.8	—	1	(FIX)	FC of ICA cmpnts	(→)	Shrinkage LDA	1081	LOOCV	75.3	47.9	85.2	—
SCZ	Kim *et al.*[98]	50 (43/7) 35.9 ± 13.6	50 (34/16) 35.5 ± 11.9	—	1	HM, WM & CSF	FC among 116 AAL ROI	(→)	DNN with L_1-norm reg	6670	5-fold CV	86	86	85	Data adopted from COBRE†
SCZ/BPD	Rashid *et al.*[80]	60 (47/13) 35.9 ± 12.0	61 (33/28) 35.4 ± 11.6	BPD 38 (18/20) 39.0 ± 10.9	1	HM & dHM	Dynamic FC among ICA-extracted networks	PCA + EM algorithm & ICA	SVM	15	10-fold CV	84	83	90	—
MDD	Lord *et al.*[99]	21 (13/8) 37.9 ± 11.4	22 (13/9) 34.6 ± 6.2	—	1	HM, GM, CSF & GS	Graph metrics for FC among 95 AAL ROI	mRMR algorithm	SVM	25	Bagging	99	99	99	—
MDD	Wei *et al.*[100]	20 (10/10) 34.3 ± 8.2	20 (14/6) 30.8 ± 8.7	—	1	—	Hurst exponents of ICA cmpnts	ICA	SVM	12	LOOCV	90	95	85	—
MDD	Cao *et al.*[101]	39 (16/23) 28.0 ± 7.5	37 (23/14) 28.2 ± 6.5	—	1	HM, GS, WM & CSF	FC among 90 AAL ROI	PDF	SVM	31/35	LOOCV	84.2	—	—	—
MDD	Guo *et al.*[102]	36 (14/22) 28.4 ± 9.7	27 (12/15) 26.8 ± 9.6	—	1	HM	Graph metrics for FC among 90 ROI	—	NN	30	(CV)	90.5	—	—	—
ADHD	Fair *et al.*[103]	80 ADHD-I 11.5 112 ADHD-C 10.3	455 14.4	—	6	GS, VT, WM, their derivatives, & HM	FC among 160 ROI	—	SVM-based MVPA	150	LOOCV	82.7 77.0	78.9 75.0	86.5 76.9	Data adopted from ADHD-200, peak performance for each ADHD type
ADHD	Sidhu *et al.*[104]	*Training data* 137 ADHD-C 98 ADHD-I *Test data* 51 ADHD-C 26 ADHD-I	423 94	—	7	—	Waveforms processed by FFT	kPCA-st	SVM	668	10-fold CV	76.0 66.7	—	—	Data adopted from ADHD-200, peak performance achieved by incorporation of phenotypic data

Table 1. (*Continued*)

Disorder	Reference	Number of participants (male/female) Mean age ± SD (if available)			Number of imaging sites	NV	Features	Dimensionality reduction & feature selection	Classifier	Number of features	Model evaluation	Classification performance			Remarks
		Patients	Controls	Other								Accuracy (%)	Sensitivity (%)	Specificity (%)	
ADHD	Dey *et al.*[105]	*Training data* 135 (Age range, 7–22)	231	—	4	WM, CSF & HM	Graph metrics for FC among 190 ROI	Thresholding (FC values)	SVM	2	LOOCV	60.6–75.6	9.5–58.8	65.8–100.0	Data adopted from ADHD-200, range of performance in four sites
		Test data 45 (Age range, 7–26)	76									48.0–82.4	20.8–50.0	62.5–92.6	
ADHD	Deshpande *et al.*[106]	173 ADHD-I 260 ADHD-C	744	—	9	WM, CSF & HM	Various connectivity metrics	Thresholding (*t*-statistics)	FCC-ANN	200	LOOCV	~90	—	—	Data adopted from ADHD-200
ASD	Anderson *et al.*[107]	40 (40/0) 22.7 ± 7.4	40 (40/0) 21.6 ± 7.4	—	1	FST, WM & CSF	FC among 7266 ROI	Thresholding (FC values)	FC-based score	~90,000	LOOCV	79	83	75	—
ASD	Nielsen *et al.*[108]	447 (396/51) 16.6 ± 8.1	517 (426/91) 16.9 ± 7.6	—	16	FST, WM, CSF & HM	FC among 7266 ROI	Thresholding (FC values)	FC-based score	(<26, 393, 745)	LOOCV	60	62	58	Data adopted from ABIDE
ASD	Uddin *et al.*[109]	20 (16/4) 10.0 ± 1.6	20 (16/4) 10.0 ± 1.6	—	1	—	ICA cmpnts	ICA	LR	1	LOOCV	78	75	80	—
ASD	Chen *et al.*[110]	126 (108/18) 17.3 ± 6.0	126 (95/31) 17.1 ± 5.7	—	13	HM, WM, VT, their derivatives & GS	FC among 220 ROI	—	RF	100	(RF)	91	89	93	Data adopted from ABIDE
ASD	Iidaka [111]	312 (273/39) 13.2 ± 3.1	328 (267/61) 12.9 ± 3.0	—	12	HM, WM & CSF	FC among 90 AAL ROI	Thresholding to Cohen's *d*	PNN	632	LOOCV	89.4	91.9	86.9	Data adopted from ABIDE
ASD	Plitt *et al.*[112]	59 17.7 ± 2.7	59 18.3 ± 3.1	—	3	HM, dHM, WM & VT	FC among ROI (three sets)	(→)	L2LR/RBF-SVM	Variable	LOOCV	76.7 / 76.7	75.0 / 70.0	78.3 / 83.3	Data adopted from ABIDE
ASD	Chen *et al.*[113]	112 (101/11) 14.8 ± 1.7	128 (104/24) 14.6 ± 1.6	—	6	HM, WM, CSF & GS	Frequency-specific FC among 142 ROI	F-score	SVM	60	LOOCV	79.2	77.8	80.5	Data adopted from ABIDE

Table 1. (*Continued*)

Disorder	Reference	Number of participants (male/female) Mean age ± SD (if available)			Number of imaging sites	NV	Features	Dimensionality reduction & feature selection	Classifier	Number of features	Model evaluation	Classification performance			Remarks
		Patients	Controls	Other								Accuracy (%)	Sensitivity (%)	Specificity (%)	
ASD	Yahata *et al.*[114]	*Training data* 74 31.4 ± 8.5 *Test data* 44 24.0 ± 6.6	117 30.4 ± 8.0 44 24.0 ± 5.0	—	3	For FC calculation: HM, GS, WM, & CSF. For feature selection: age, sex, eye status & medication status	FC among 140 ROI	L_1-SCCA	SLR	16	LOOCV	85	80	89	(See also Fig. 2)
				—	7							75	75	75	

[†]COBRE, http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html.

AAL, automated anatomical labeling; ABIDE, Autism Brain Imaging Data Exchange[115]; ADHD, attention-deficit hyperactivity disorder; ADHD-C, ADHD combined subtype; ADHD-I, ADHD inattentive subtype; ADHD-200, ADHD-200 Global Competition[116]; ALFF, amplitude of low frequency fluctuations; ASD, autism spectrum disorder; cmpnts, components; BC, betweenness centrality; BDopt, block diagonal optimization; BPD, bipolar disorder; CN, complex network (graph analysis); COBRE, Center for Biomedical Research Excellence; CSF, cerebrospinal fluid; CV, cross validation; dHM, derivative of head motion; DNN, deep neural network; EM, expectation maximization; FC, functional connection; FCC-ANN, fully connected cascade artificial neural network; FFT, fast Fourier transform; FIX, noise components removal by FIX algorithm[117]; FST, facial soft tissue; GM, gray matter; GS, global signal; HM, head motion; HN, handedness; ICA, independent component analysis; IDA, intrinsic discriminative analysis; KNN, k-nearest neighbors; kPCA-st, kernelized principal component analysis over space and time; LASSO, least absolute shrinkage and selection operator; LDA, linear discriminant analysis; LLE, locally linear embedding; LOOCV, leave-one-out cross-validation; LR, logistic regression; L2LR, L2-regularized LR; MDD, major depressive disorder; MDL, minimum description length; MIC, maximal information coefficient; mRMR, minimum redundancy maximum relevance; MVPA, multi variate pattern analysis; NN, neural network; NV, nuisance variables; PCA, principal component analysis; PDF, probability density function; PNN, probabilistic neural network; RBF, radial basis function (kernel); RC, rank correlation; ReHo, regional homogeneity; RF, random forest; reg, regularization; ROI, region of interest; SCCA, sparse canonical correlation analysis; SCZ, schizophrenia; SD, standard deviation; SLR, sparse logistic regression; SVM, support vector machine; uSIB, unaffected siblings; V-ELM, voting-based extreme learning machine; VT, ventricles; WM, white matter.

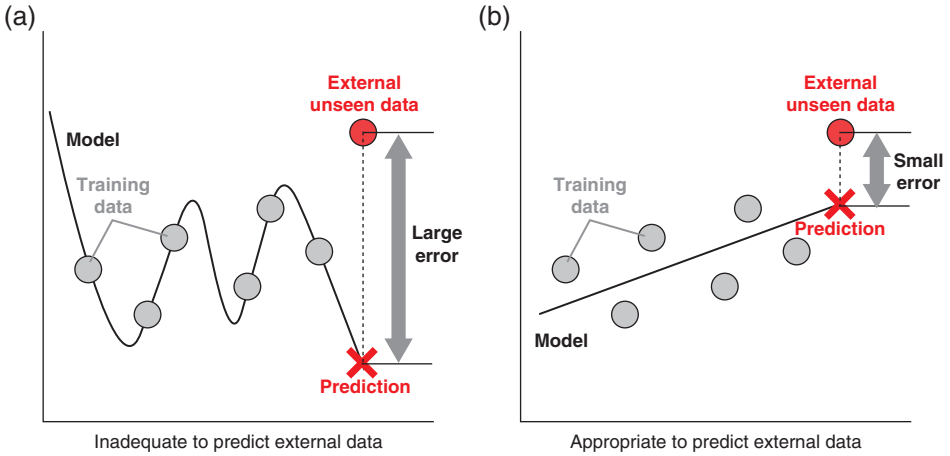

Figure 1. Schematic illustrating overfitting. (a) When the parameter dimension of a model exceeds the sample size of the training data (discovery cohort), the model fits to not only the true underlying relationship but also the noise structure inherent in the data. In this case, the ability of the model to predict external unseen data (validation cohort) becomes extremely poor. (b) An appropriate procedure to circumvent overfitting is to balance model complexity and the sample size of the training data. In addition to simply increasing the sample size, this may be achieved by supervised (such as least absolute shrinkage and selection operator [LASSO] and Elastic Net) and unsupervised (e.g., principal component and independent component analyses) techniques for dimensionality reduction and feature selection.

not replicated beyond the training data, therefore overfitting will not allow for accurate predictions of independent data. Overfitting occurs when the complexity of the model, for example, the number of explanatory variables (p), exceeds the sample size (n) (Fig. 1a). Excessive model complexity results in an unintended fit into the noise structure of the data, as is often the case with neuroimaging studies because the dimension of model parameters tends to exceed the number of participants by a large factor ($p \gg n$). For example, the number of voxels is typically $\sim 10^5$ for functional and diffusion tensor images and $\sim 10^6$ for structural MR images, whereas the number of participants incorporated into a single study is typically on the order of 10^2 to 10^3. Balancing the complexity of a model against the available sample size of a training dataset (i.e., increasing n and reducing p) is therefore essential to improving prediction accuracy for independent data (Fig. 1b).

Increasing the sample size to the order of $\sim 10^3$ is not unrealistic if the design of a study allows for the incorporation of data from previous and ongoing large-scale imaging campaigns (e.g., imaging campaigns for ADHD,[116] ASD,[115] and the general population[125]). Anonymized imaging and demographic data from these and other imaging projects are available at open-data repositories.[126,127] Another more effective approach to mitigate the overfitting problem is to reduce the model's dimension, which is achieved via an unsupervised method for dimensionality reduction and/or a supervised method for feature reduction.[128] Using this approach is particularly important when the classification is based on functional connectivity, where the parameter dimension can be as large as the number of imaging elements squared. An unsupervised method for dimensionality reduction is typically employed during image data preprocessing. One simple and popular method is to segment the brain into subregions (typically $\sim 10^2$ regions) using anatomical atlases, and then extract a representative measurement from each segmented region. Performing principal component[129] and independent component analyses[120] on fMRI time-series data and computing the mutual correlation of the extracted components is another popular unsupervised method. A supervised method of dimensionality reduction is achieved through a regularization procedure that is built into machine-learning algorithms, such as LASSO and Elastic Net techniques. In these cases, a loss function that is used to optimize the model incorporates a penalty term that controls the complexity of the model. In LASSO and Elastic Net, the penalty term includes a hyperparameter that specifies the sparsity of model coefficients (sparsity meaning fewer non-zero

coefficients and therefore fewer parameters in the model) that is determined by a cross-validation procedure.[128] Some previous studies have used the regularization technique to reduce the number of features.[94,98,114]

Proper handling of NV

A particular pattern of interaction between disorder-specific and non-specific features (i.e., NV) can be exploited to inflate the classification performance. A simple but illustrative example is where data from one site contains younger patients and data from another site contains older control patients. In this case, age can be used to predict this particular dataset, but not the general population. Disease-non-specific features include both the demographic factors of participants, such as age,[130] sex,[131] and medication profiles,[132] as well as conditions related to data acquisition, such as scanner specifications,[133] imaging parameters,[134] and instructions to participants.[135] Previous studies have indicated that all of these features affect MR data interpretation. Again, the pattern of interaction is often unique to the training data so that the classification scheme exploiting this interaction structure is doomed to yield catastrophic predictions for independent data.

Construction of a generalizable classifier

To avoid overfitting and unintended exploitation of disease-non-specific characteristics in the training data, it is essential that imaging studies proceed with a large sample of data across multiple imaging sites[136,137] and with appropriate algorithms that both balance the complexity of the model and sample size and handle the effects of NV in selecting disease-specific features from the neuroimaging data. In Table 1, there are only three studies[104,105,114] wherein development of the classifier involved both training dataset (discovery cohort) and test dataset (validation cohort) from multiple imaging sites. However, in two[104,105] of these studies, the training and test datasets were not completely independent because they were derived from the almost identical set of imaging sites. In one study,[104] classification performance for the test dataset was not better than that of a classifier based only on disease-non-specific, phenotypic information. In the other study,[105] classification performance was characterized by unbalanced sensitivity (59% at most) and specificity (up to 90–100%) in both the training and test datasets. The generalization of performance to independent validation cohorts was untested in these studies. In contrast, the remaining study[114] formed a classifier based on the dataset acquired at three sites. The construction of the classifier involved an L_1-norm regularized sparse canonical correlation analysis (L_1-SCCA)[138] and sparse logistic regression (SLR)[139] for concurrent removal of NV effects and dimensionality reduction (Fig. 2). To date, this is the only study in which high classification performance was generalized to an independent cohort. Again, this underscores the importance of appropriate handling of overfitting and NV. For reference, if one extends the scope of the survey to include previous structural MRI studies, two studies[140,141] have reported the construction of classifiers for patients with schizophrenia and normal controls and replicated classification performance in independent samples.

Utility of neuroimaging-based classifiers in psychiatry

What is the utility of generalization-proof neuroimaging-based classifiers in psychiatry? Can such a system be used in an outpatient psychiatry clinic as a screening tool to identify those in need of further medical assessment, or can it be used to provide complementary information for those patients whose diagnoses are hard to determine using the standard diagnostic procedure? Realistically, the utility of such a system as a screening tool must be considered at the intersection of the prevalence rate of a disorder, classification sensitivity and specificity, and scanning cost. For example, screening individuals for a disorder with a 1% prevalence rate (e.g., ASD) using a classifier with a sensitivity and specificity of 80% detects one true-positive case for approximately 25 false-positive cases. Whether or not this is acceptable from the viewpoint of medical economy actually depends on the efficiency and reliability of other existing screening tools. In this sense, a classifier may better be used in such a selected case where existing diagnostic procedures do not yield a definitive diagnosis and complementary information is needed to narrow down the possibilities.

Yet another *raison d'être* of a classifier is to aid in the recasting of mental disorders into a biology-oriented stratification instead of a symptom-based stratification (Fig. 3). A classifier quantitatively delineates one's internal state as a likelihood of a

specific disorder. A set of classifiers, each constructed for a certain type of disorder, will then provide a set of axes in a multi-dimensional space of mental disorders, where each individual is represented as a point, and a group of individuals with one type of disorder is represented as a cloud of points. Investigating the spatial distribution of clouds of patients with specific disorders will enable the quantitative characterization of mutual relationships for disorders, providing important clues for revising the definitions of disorders toward a biological viewpoint. For example, determining new axes by maximizing variance will generate new categorizations of disorders.

A recent investigation[114] used a classifier for ASD to investigate its relation with other disorders, such as schizophrenia, MDD, and ADHD. The pattern of functional connectivity that is selected in a classifier, namely, the weighted linear summation (WLS) of the correlation indices of connections, represents the 'ASD-ness' of each individual. This biological measure was calculated for a group of patients with

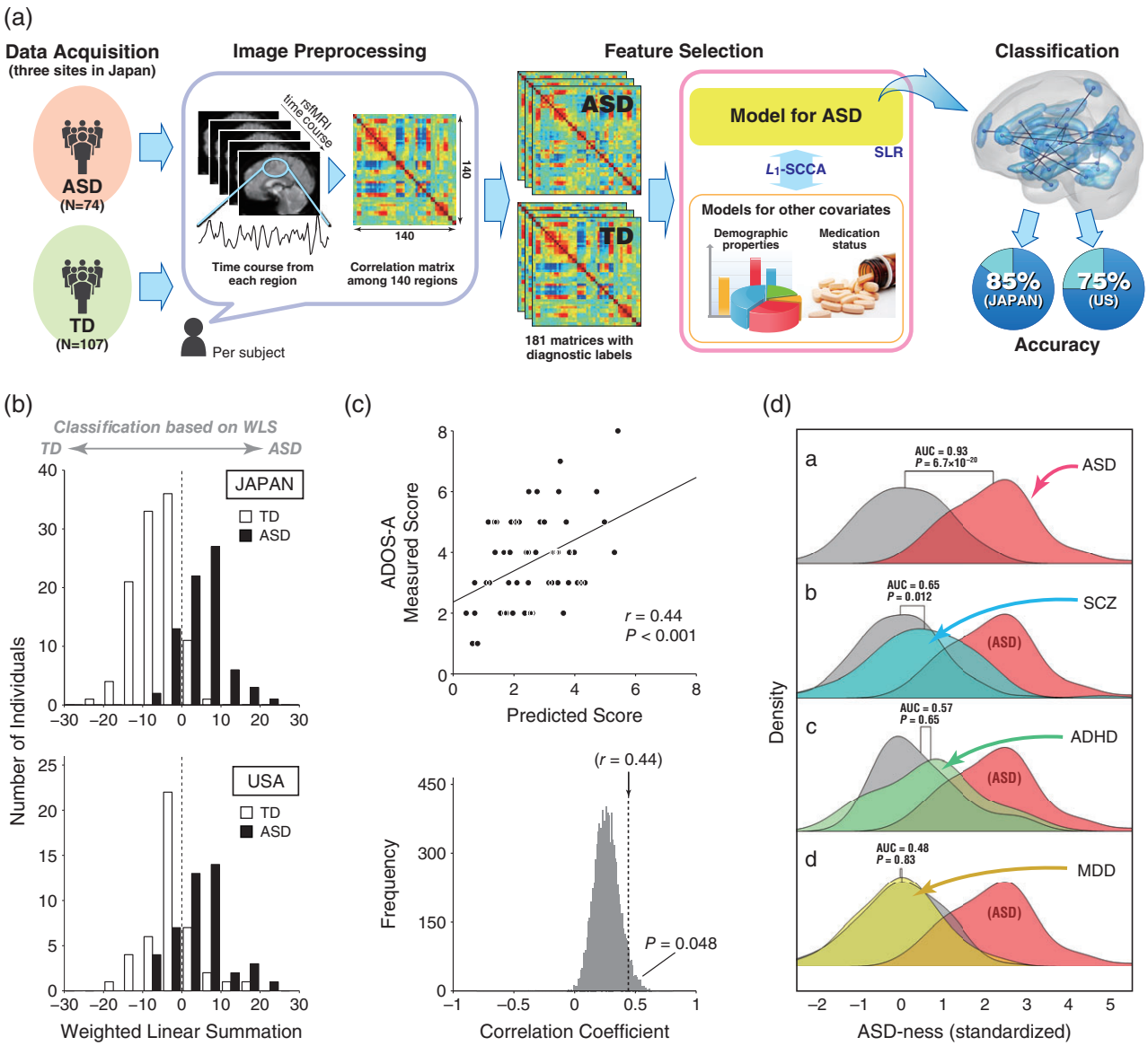

Figure 2. Legend on next page.

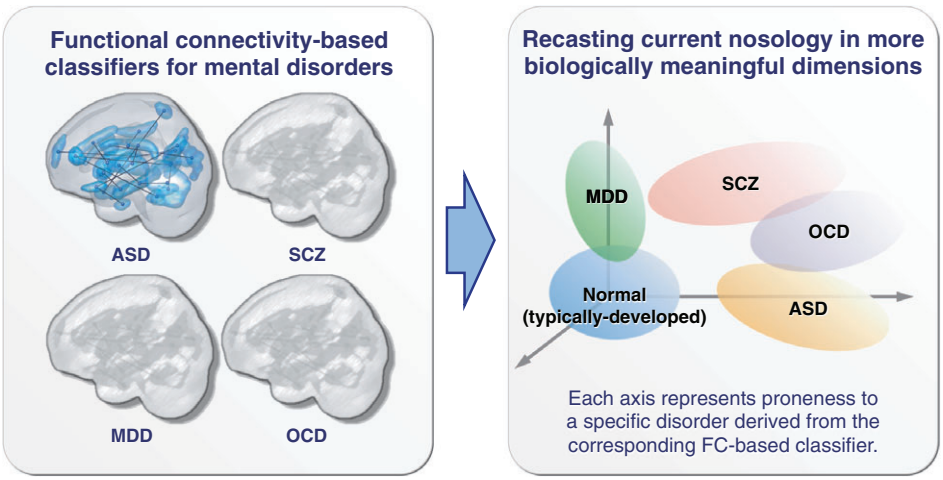

Figure 3. Schematic illustrating the redefinition of mental disorders using functional connectivity (FC)-based classifiers. An FC-based classifier represents one's own proneness to a specific disorder. A set of classifiers, each constructed for a certain type of disorder, will then form a multi-dimensional fiducial space of mental disorders, in which mutual relations between the disorders can be quantitatively examined. ASD, autism spectrum disorder; MDD, major depressive disorder; OCD, obsessive–compulsive disorder; SCZ, schizophrenia.

schizophrenia, MDD, or ADHD, and their controls. The results of this study revealed that only in the case of schizophrenia did the ASD-ness distribution of patients and controls differ significantly, suggesting that ASD and schizophrenia share some intrinsic functional connections. Interestingly, this finding was consistent with previous behavioral[142,143] and genetic studies[144,145] (Fig. 2).

Prediction of treatment response and prognosis

Predicting clinical sequelae in a patient is another important area in psychiatry research to which machine-learning techniques may be applied. The features revealed as important for prediction may play a pivotal role in establishing the long-awaited

Figure 2. An example of development of a generalizable classifier for autism spectrum disorder (ASD).[114] (a) Yahata *et al.*[114] recently developed a generalizable classifier for ASD by applying a novel machine learning technique to a set of multi-site, resting-state functional connectivity magnetic resonance imaging (rs-fcMRI) data acquired from 74 adults with ASD and 107 demographically matched, typically developed (TD) individuals. For each individual, a pairwise interregional correlation matrix was formed for 140 regions of interest (ROI) using the mean regional time course of blood oxygenation level-dependent (BOLD) signals calculated from the preprocessed images. The set of correlation matrices, as well as associated demographic information (diagnostic label, age, sex, site scanned, medication status, etc.), was submitted to a cascade of two machine-learning techniques: L_1-norm regularized sparse canonical correlation analysis (L_1-SCCA)[138] and sparse logistic regression (SLR).[139] Briefly, the L_1-SCCA procedure selects those functional connections (FCs) that are correlated with the individuals' diagnostic labels (ASD or TD) while discarding the FC under the influence of NV. Then, the remaining FCs (~10% of the total FCs considered initially) were further submitted to the SLR procedure, and those FCs most relevant to the diagnostic label were selected. (b) The result of these statistical procedures is a set of weights for the 16 finally selected FCs (0.2% of the total FCs). The weighted linear summation (WLS) of the corresponding correlation indices plus constant is the argument for the logistic function that maps the individual's WLS to the prediction of a diagnosis. The classification accuracy for the Japanese training data was 85%, with an area under the curve (AUC) of 0.93, indicating highly reliable classification. This classifier was further applied to an independent dataset adopted from the US ABIDE project. We confirmed that the reliability of the classifier was generalized to this independent cohort with a diverse ethnic background, in that the classification accuracy was 75% (AUC = 0.76), and the probability of achieving this accuracy was as small as $P = 1.4 \times 10^{-6}$. (c) In addition, the 16 FCs predicted the socio-communicative behaviors of ASD individuals were measured by the Autism Diagnostic Observation Schedule (ADOS). (d) Finally, using this classifier, we measured the 'ASD-ness' of patients with schizophrenia, MDD, or ADHD and their healthy controls. These results demonstrate that the ASD-ness quantified the spectrum of the four disorders as follows: schizophrenia was close to ASD, ADHD was distant from ASD, and MDD was farthest from ASD.

goal of 'precision medicine' in psychiatry[146]: providing a quantitative means to optimize therapeutic strategies on an individual basis. Individually optimized treatment may lead to reductions in the period of patient disability, and a decrease in adverse outcomes; both of these results present a significant benefit to public health. Table 2 lists a selection of previous work that has employed machine-learning classification to predict both treatment response[147–155] and prognosis[156–160] in psychiatric patients. For treatment response, the target of prediction includes both pharmacological (including antipsychotics for schizophrenia,[148] ω-3 fatty acids for ultra-high risk for psychosis,[150] antidepressants for MDD[155] and obsessive–compulsive disorder,[154] and methylphenidate for ADHD[151]) and psychological intervention (cognitive behavioral therapy for alcohol-dependent individuals[147] and individuals with social anxiety disorder[152]). For prognosis, the target of prediction includes state transition from at-risk mental state to psychosis,[156] chronic course of MDD,[159] resilience to or development of post-traumatic stress disorder symptoms,[157,158] and relapses in alcohol dependence.[160] Classifier development typically proceeds with one or a few predefined sets of demographic, clinical, and biological parameters of individuals. As with the prediction of diagnosis, a support vector machine is the most popular classification algorithm. The accuracy of the predictions is generally greater than 70% or 0.7 represented as the area under the curve (AUC). However, generalizations of this performance in an independent cohort remained untested as yet. Except for a few studies, the sample size tends to be small to modest ($n < 100$). The use of neuroimaging, often combined with demographic and clinical measures, has been increasing recently, which should provide important information for understanding the neural correlates of treatment response and state transition. Because the prediction of prognosis entails temporal changes within an individual, future work may need to address the pattern of interaction between the features used in the classification and the environmental factors affecting an individual in order to improve prediction accuracy.

THERAPEUTIC APPLICATION

Neurofeedback, a descendant of the more general biofeedback, is a non-invasive method for modulating the activity of the whole brain or specific regions by providing an individual with feedback related to their current brain state. Upon presentation of this feedback, a person may attempt to self-regulate (increase or decrease) their activation levels towards a state appropriate for bringing about a desired behavior.[162] The successful modulation of activity is the result of operant conditioning, which can be therapeutic if employed in a clinical setting. The first wave of feedback studies came in the second half of the last century, when biofeedback with electroencephalography (EEG) was applied to a wide range of mental disorders, including ADHD, mood disorders, anxiety disorders, obsessive–compulsive disorder, alcoholism, and numerous others. However, no robust evidence yet exists to strongly support the clinical efficacy of EEG-based biofeedback.[162,163]

The second wave of feedback studies arrived at the turn of this century, when real-time fMRI was introduced as a new tool for neurofeedback (hereafter 'rtfMRI-nf'). Thanks to its high spatial resolution (~2–3 mm) and to the high computing power that supports online image reconstruction (with a delay of less than 1–2 s), rtfMRI-nf allows focal modulation of the brain with a latency of only 2–3 s. The regions of the brain that have been targets of modulation in previous studies include the insula,[164,165] amygdala,[166,167] prefrontal cortex,[168] and anterior cingulate cortex,[165] among others. In some studies, participants were reported to be able to modulate the activity of the target region; however, the lack of appropriate control conditions generally precluded a definitive interpretation that neurofeedback was indeed the primary factor that accounted for the observed modulation.[162,163]

More recently, interregional functional connectivity has become a target for neurofeedback studies. In one study, healthy individuals underwent 4 days of rtfMRI-nf sessions to train them to modulate the functional connection between two designated regions. Participants successfully changed their resting-state functional connectivity between the two regions and this alteration remained for more than 2 months, an encouraging result in terms of the therapeutic application of rtfMRI-nf.[169] For example, the set of functional connections shown in Figure 2 may be used as a possible therapeutic target for individuals with ASD during the rtfMRI-nf session. The WLS of the ASD classifier could be estimated for the participant in the scanner during an rtfMRI-nf

Table 2. Summary of previous developments in machine-learning-based classifiers for predicting patient (a) treatment response and (b) prognosis

References	Target of prediction	Participants	Classifier	Features			Model evaluation	Performance of prediction	
				Used for classification	Neuroimaging included?	Number of Features		Accuracy (%)	AUC
(a) Treatment response									
Connor *et al.*[147]	Abstinence at 12 weeks of CBT for alcohol-dependent subjects	[CBT group] 60 (13) alcohol-dependent individuals in training (test) dataset [CBT + medication group] 53 (13) alcohol-dependent individuals in training (test) dataset	(1) DA, (2) DT, (3) BN	Demographics, alcohol dependence history, etc.	No	Variable	—	(1) 31,† (2) 77,† (3) 77†	—
Nejad *et al.*[148]	Improvement on negative symptom after 7 months of quetiapine monotherapy	14 antipsychotic-naïve, first-episode patients with schizophrenia	SVM	Network components modulated by an n-back working memory task	Yes	3	LOOCV	78.6	—
Perlis[149]	Response to first or second pharmacological treatment trial in STAR*D study	[Training data] 538 MDD patients not remitted and 1033 MDD patients remitted [Test data] 199 MDD patients not remitted and 324 MDD patients remitted [Validation data] 147 MDD patients not remitted and 314 MDD patients remitted	(1) LR, (2) NB, (3) RF, (4) SVM	Demographics, QIDS scores, psychiatric comorbidity, illness course features	No	15	10-fold CV	(1) 69.1 (2) 69.2, (3) 69.8, (4) 67.7 (1) 66.5, (2) 67.1, (3) 67.9, (4) 66.2 (1) 70.3	(1) 0.714, (2) 0.716, (3) 0.706, (4) 0.697 (1) 0.712, (2) 0.698, (3) 0.693, (4) 0.706 (1) 0.719
Amminger *et al.*[150]	Response to 12 weeks of long-chain ω-3 fatty acids treatment	[ω-3 group] 40 individuals at UHR of psychosis (22 responders) [Placebo group] 40 individuals at UHR of psychosis (12 responders)	GPC	Fatty acid compositions measured at baseline	No	21	LOOCV	86.7 79.2	—
Kim *et al.*[151]	Response to 8-week open-label trial of methylphenidate	83 children with ADHD (48 responders)	(1) SVM, (2) DT, (3) RF, (4) LRR	Demographic, clinical, genetic, environmental and neuroimaging measures	Yes	Variable	10-fold CV	(1) 84.6, (2) 69.2, (3) 73.1, (4) 76.9	(1) 0.84, (2) 0.61, (3) 0.79, (4) 0.73
Månsson *et al.*[152]	Long-term treatment response (at 1-year follow-up) to cognitive behavioral therapy	26 participants with a primary diagnosis of social anxiety disorder (12 responders)	SVM	Regional BOLD responses to self-referential criticism paradigm	Yes	2	LOOCV	91.7	0.89
Patel *et al.*[153]	Response to 12 weeks of open-label pharmacological treatment	19 patients with late-life depression (9 responders)	ADTree	DTI number of tracks from aSN and rsfMRI functional connectivity index from dDMN	Yes	2	LOOCV (nested)	89.5	—

Table 2. (*Continued*)

References	Target of prediction	Participants	Classifier	Features			Model evaluation	Performance of prediction	
				Used for classification	Neuroimaging included?	Number of Features		Accuracy (%)	AUC
Yun *et al.*[154]	Response to 4 months of serotonin reuptake inhibitors treatment	56 unmedicated patients with OCD (25 responders) and 75 healthy controls	SVM	Cortical surface-area-based and thickness-based individualized structural covariances	Yes	12	Permutation	89.0	—
Iniesta *et al.*[155]	Remission after 12-week trial of treatment with antidepressants (escitalopram and nortripyline)	[Allocated group] 793 individuals with MDD allocated to either escitalopram (n = 465) or nortriptyline (n = 328) arm	Elastic net	Demographics, severity measures of MDD, index for stressful life events, medication history	No	41	10-fold CV	—	0.72
		[Randomly-allocated group] 450 individuals with MDD allocated randomly to two drug arms				48			0.74
(b) Prognosis									
Koutsouleris *et al.*[156]	Transition to psychosis over 4 years of follow-up	33 individuals in an at-risk mental state of psychosis with completed follow-up (15 with TP and 18 without TP) and 17 matched HC	SVM	RAVENS map calculated from the structural MRI images	Yes	17	5-fold CV	82[‡]	—
Galatzer-Levy *et al.*[157]	Non-remission of PTSD in trauma survivors based on their early trauma responses	957 subjects admitted to emergency following traumatic events (17% non-remitting)	SVM	Demographics, physiological responses, medications, etc.	No	16	10-fold CV	—	0.82
Karstoft *et al.*[158]	Resilience or post-traumatic stress development over 2.5 years after deployment	561 Danish soldiers deployed to Afghanistan in 2009	SVM	Demographics, history of military service, psychometric measures related to PTSD	No	6 / 9	10-fold CV	—	0.84 / 0.88
Schmaal *et al.*[159]	Chronic course of MDD over 2 years of follow-up	118 individuals with MDD (23 chronic MDD)	GPC	(1) Activation pattern in fMRI during faces task (2) clinical characteristics	Yes	(1) 5, (2) 9	LOOCV	(1) 73, (2) 69	—
Seo *et al.*[160]	Relapse of alcohol-dependent patients over 3-month assessment period	46 abstinent alcohol-dependent patients (30 relapsers)	(1) NB, (2) SVM, (3) RSLVQ	ROI-basis features in structural and functional MRI	Yes	48	LOOCV	(1) 72.7, (2) 74.8, (3) 79.4	—

[†]Evaluated on the respective test dataset.
[‡]Three-group classification performance.
ADHD, attention-deficit hyperactivity disorder; ADTree, alternating decision tree; aSN, anterior salience network; BN, Bayesian network; BOLD, blood oxygenation level-dependent; CBT, cognitive behavioral therapy; CV, cross validation; DA, discriminant analysis; dDMN, dorsal default mode network; DT, decision tree; DTI, diffusion tensor imaging; GPC, Gaussian process classification; HC, healthy control; LOOCV, leave-one-out cross validation; LR, logistic regression; LRR, logistic ridge regression; MDD, major depressive disorder; MRI, magnetic resonance imaging; NB, Naïve Bayes; OCD, obsessive compulsive disorder; PTSD, post-traumatic stress disorder; QIDS, Quick Inventory of Depressive Symptoms; RAVENS, regional analysis of volumes in normalized space; RF, random forest; ROI, region of interest; rsfMRI, resting-state functional magnetic resonance imaging; RSLVQ, robust soft learning vector quantization; STAR*D, sequenced treatment alternatives to relieve depression[161]; SVM, support vector machine; TP, transition to psychosis; UHR, ultra high-risk.

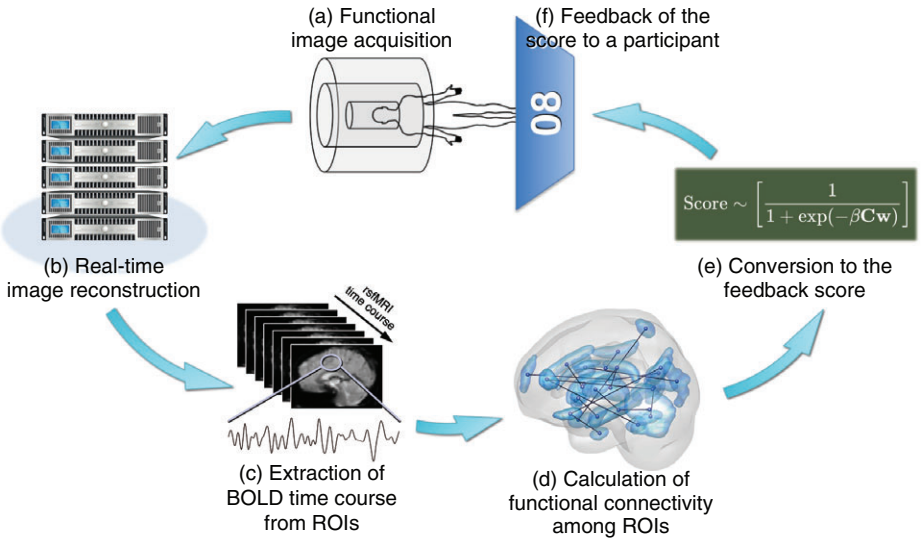

Figure 4. Schematic illustrating real-time neurofeedback. (a) A participant's whole-brain activity is recorded by functional magnetic resonance imaging (fMRI) at intervals of 2–3 s. (b) The acquired image is reconstructed within a short delay, typically 1–2 s. (c) The image is transferred to a local computer and processed, so that representative blood oxygenation level-dependent (BOLD) time series are extracted from the target regions of interest (ROI). (d) The values of disorder-specific functional connections (e.g., those incorporated into an functional-connectivity-based classifier) are calculated and are then (e) converted to a single number (score) indicating the participant's current proneness to a disorder before being (f) fed back to a participant who then attempts to self-regulate his/her activation level toward a preferred state.

session. Thereafter, a sign-inverted WLS value could be presented to the individual with ASD as a neurofeedback target for increasing in an RL paradigm. Increases in the neurofeedback score lead to reductions in the WLS value. Successful completion of the rtfMRI-nf sessions might then bring about normal functional connectivity dynamics in the participant (Fig. 4). Although it still sounds surreal, efforts in the development of functional connectivity-based rtfMRI-nf are ongoing worldwide, and may one day represent a novel therapeutic option in clinical psychiatry.

CONCLUSIONS

Here, we have reviewed previous applications of computational neuroscience in psychiatry research and have thereby discussed its potential role in achieving a better understanding of the etiology and pathophysiology of mental disorders. While theory-driven approaches will provide an integrated view of the mental disorders by connecting the disorder-specific features derived at each level of the brain's organization, data-driven approaches may provide a common platform for making biology-based, quantitative diagnoses of, and treatments for, mental disorders. In fact, the application of machine-learning techniques to resting-state functional connectivity holds great promise as a means to enabling simultaneous diagnosis and treatment, thereby establishing 'theranostics'[170,171] for the first time in clinical neuropsychiatry in the near future.

ACKNOWLEDGMENT

This work was conducted under the 'Development of BMI Technologies for Clinical Application' of the Strategic Research Program for Brain Sciences supported by the Japan Agency for Medical Research and Development (AMED). This work was partially supported by the Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) from AMED and the Japan Society for the Promotion of Science (JSPS) KAKENHI (16K10233 to N.Y.). We thank Drs Ichio Aoki, Takashi Itahashi, and Kentaro Morita for their technical assistance in preparing the manuscript.

DISCLOSURE STATEMENT

M.K., N.Y., and K.K. are inventors of a patent owned by Advanced Telecommunications Research (ATR) Institute International related to the present work

(PCT/JP2014/061543 [WO2014178322] and PCT/JP2014/061544 [WO2014178323]). M.K. and N.Y. are inventors of patent applications submitted by ATR Institute International related to the present work (JP2015-228970) and (JP2015-005346). M.K. is an inventor of patent applications submitted by ATR Institute International related to the present work (JP2014-262673), (PCT/JP2012/078136 [WO2013069517]), (JP2013-259554), (JP2013-122427), and (JP2011-244048). For the past 3 years, K.K. declares the following potential conflicts of interest (although they are all unrelated to the current study): K.K. has received research grants from Astellas Pharma, GlaxoSmithKline, Dainippon-Sumitomo, Eisai, MSD, and Yoshitomi; K.K. has received honoraria for lectures by Daiichi-Sankyo, Otsuka, Meiji Seika, MSD, Astellas Pharma, Yoshitomi, Novartis, Eli Lilly, Dainippon-Sumitomo, Janssen, GlaxoSmithKline, and Pfizer. Also, M.K. has received an honorarium for a lecture by the 5th Annual Computational Neuroscience Meeting.

AUTHOR CONTRIBUTIONS

N.Y., K.K., and M.K. wrote the manuscript and checked the overall consistency.

REFERENCES

1. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol. Psychiatry* 2012; **17**: 1174–1179.
2. Owen MJ. New approaches to psychiatric diagnostic classification. *Neuron* 2014; **84**: 564–571.
3. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ. Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study. *Lancet* 2013; **382**: 1575–1586.
4. Bloom DE, Cafiero E, Jané-Llopis E *et al. The Global Economic Burden of Noncommunicable Diseases*. World Economic Forum, Geneva, 2012.
5. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Association, Washington, DC, 2013.
6. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. World Health Organization, Geneva, 1992.
7. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Med.* 2013; **11**: 126.
8. Adam D. Mental health: On the spectrum. *Nature* 2013; **496**: 416–418.
9. Cuthbert BN, Insel TR. Toward new approaches to psychotic disorders: The NIMH Research Domain Criteria project. *Schizophr. Bull.* 2010; **36**: 1061–1062.
10. Morris SE, Cuthbert BN. Research domain criteria: Cognitive systems, neural circuits, and dimensions of behavior. *Dialogues Clin. Neurosci.* 2012; **14**: 29–37.
11. Adams RA, Huys QJM, Roiser JP. Computational psychiatry: Towards a mathematically informed understanding of mental illness. *J. Neurol. Neurosurg. Psychiatry* 2016; **87**: 53–63.
12. Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* 2016; **19**: 404–413.
13. Sutton RS, Barto AG. *Reinforcement Learning*. MIT Press, Cambridge, MA, 1998.
14. Camerer CF. *Behavioral Game Theory*. Princeton University Press, Princeton, NJ, 2011.
15. Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* 2011; **14**: 154–162.
16. Kishida KT, King-Casas B, Montague PR. Neuroeconomic approaches to mental disorders. *Neuron* 2010; **67**: 543–554.
17. Castellanos FX, Di Martino A, Craddock RC, Mehta AD, Milham MP. Clinical applications of the functional connectome. *Neuroimage* 2013; **80**: 527–540.
18. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage* 2017; **145**: 137–165.
19. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends Cogn. Sci.* 2012; **16**: 72–80.
20. Rumelhart DE, McClelland JL, PDP Research Group (eds). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. *Psychological and Biological Models*, *Vol. 2*. MIT Press, Cambridge, MA, 1986.
21. McClelland JL. The organization of memory. A parallel distributed processing perspective. *Rev. Neurol. (Paris)* 1994; **150**: 570–579.
22. Rumelhart DE, McClelland JL. On learning the past tenses of English verbs. In: Rumelhart DE, McClelland JL, PDP Research Group (eds). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. *Psychological and Biological Models*, *Vol. 2*. MIT Press, Cambridge, MA, 1986; 216–271.
23. McClelland JL, Elman J. Interactive processes in speech perception: The TRACE model. In: Rumelhart DE, McClelland JL, PDP Research Group (eds). *Parallel Distributed Processing Explorations in the Microstructure of Cognition*. *Psychological and Biological Models*, *Vol. 2*. MIT Press, Cambridge, MA, 1986; 58–121.
24. Peterson BS, Skudlarski P, Gatenby JC, Zhang H, Anderson AW, Gore JC. An fMRI study of stroop word-color interference: Evidence for cingulate subregions

subserving multiple distributed attentional systems. *Biol. Psychiatry* 1999; **45**: 1237–1258.
25. O'Brien G, Opie J. A connectionist theory of phenomenal experience. *Behav. Brain Sci.* 1999; **22**: 127–148.
26. Hoffman RE, McGlashan TH. Parallel distributed processing and the emergence of schizophrenic symptoms. *Schizophr. Bull.* 1993; **19**: 119–140.
27. Hoffman RE, Dobscha SK. Cortical pruning and the development of schizophrenia: A computer model. *Schizophr. Bull.* 1989; **15**: 477–490.
28. Cohen JD, Braver TS, O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 1996; **351**: 1515–1527.
29. Murray GK, Corlett PR, Clark L *et al.* Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol. Psychiatry* 2007; **13**: 267–276.
30. Corlett PR, Murray GK, Honey GD *et al.* Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain* 2007; **130**: 2387–2400.
31. Romaniuk L, Honey GD, King JRL *et al.* Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. *Arch. Gen. Psychiatry* 2010; **67**: 1246–1254.
32. Schlagenhauf F, Sterzer P, Schmack K *et al.* Reward feedback alterations in unmedicated schizophrenia patients: Relevance for delusions. *Biol. Psychiatry* 2009; **65**: 1032–1039.
33. Kapur S. Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* 2003; **160**: 13–23.
34. Waltz JA, Schweitzer JB, Gold JM *et al.* Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology* 2008; **34**: 1567–1577.
35. Gradin VB, Kumar P, Waiter G *et al.* Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* 2011; **134**: 1751–1764.
36. Gold JM, Waltz JA, Matveeva TM *et al.* Negative symptoms and the failure to represent the expected reward value of actions: Behavioral and computational modeling evidence. *Arch. Gen. Psychiatry* 2012; **69**: 129–138.
37. Gold JM, Strauss GP, Waltz JA, Robinson BM, Brown JK, Frank MJ. Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biol. Psychiatry* 2013; **74**: 130–136.
38. Huys QJ, Pizzagalli DA, Bogdan R, Dayan P. Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biol. Mood Anxiety Disord.* 2013; **3**: 12.
39. Sonuga-Barke E. The dual pathway model of AD/HD: An elaboration of neuro-developmental characteristics. *Neurosci. Biobehav. Rev.* 2003; **27**: 593–604.
40. Smith AJ, Becker S, Kapur S. A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. *Neural Comput.* 2005; **17**: 361–395.
41. Deserno L, Boehme R, Heinz A, Schlagenhauf F. Reinforcement learning and dopamine in schizophrenia: Dimensions of symptoms or specific features of a disease group? *Front. Psychiatry* 2013; **4**: 172.
42. Chen C, Takahashi T, Nakagawa S, Inoue T, Kusumi I. Reinforcement learning in depression: A review of computational research. *Neurosci. Biobehav. Rev.* 2015; **55**: 247–267.
43. Fett AKJ, Shergill SS, Krabbendam L. Social neuroscience in psychiatry: Unravelling the neural mechanisms of social dysfunction. *Psychol. Med.* 2014; **45**: 1145–1165.
44. Sanfey AG. Social decision-making: Insights from game theory and neuroscience. *Neuroscience* 2007; **318**: 598–602.
45. Yoshida W, Seymour B, Friston KJ, Dolan RJ. Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* 2010; **30**: 10744–10751.
46. Yoshida W, Dziobek I, Kliemann D, Heekeren HR, Friston KJ, Dolan RJ. Cooperation and heterogeneity of the autistic mind. *J. Neurosci.* 2010; **30**: 8815–8818.
47. Fett AKJ, Shergill SS, Joyce DW *et al.* To trust or not to trust: The dynamics of social interaction in psychosis. *Brain* 2012; **135**: 976–984.
48. Ong DC, Zaki J, Gruber J. Increased cooperative behavior across remitted bipolar I disorder and major depression: Insights utilizing a behavioral economic trust game. *J. Abnorm. Psychol.* 2017; **126**: 1–7.
49. Chiu PH, Kayali MA, Kishida KT *et al.* Self responses along cingulate cortex reveal quantitative neural phenotype for high-functioning autism. *Neuron* 2008; **57**: 463–473.
50. Lis S, Baer N, Franzen N *et al.* Social interaction behavior in ADHD in adults in a virtual trust game. *J. Atten. Disord.* 2016; **20**: 335–345.
51. Unoka Z, Seres I, Aspán N, Bódi N, Kéri S. Trust game reveals restricted interpersonal transactions in patients with borderline personality disorder. *J. Pers. Disord.* 2009; **23**: 399–409.
52. Koshelev M, Lohrenz T, Vannucci M, Montague PR. Biosensor approach to psychopathology classification. *PLoS Comput. Biol.* 2010; **6**: e1000966.
53. Honea R, Crow TJ, Passingham D, Mackay CE. Regional deficits in brain volume in schizophrenia: A meta-analysis of voxel-based morphometry studies. *Am. J. Psychiatry* 2005; **162**: 2233–2245.
54. Lai C-H. Gray matter volume in major depressive disorder: A meta-analysis of voxel-based morphometry studies. *Psychiatry Res.* 2013; **211**: 37–46.
55. Selvaraj S, Arnone D, Job D *et al.* Grey matter differences in bipolar disorder: A meta-analysis of voxel-based morphometry studies. *Bipolar Disord.* 2012; **14**: 135–145.

56. Stanfield AC, McIntosh AM, Spencer MD, Philip R, Gaur S, Lawrie SM. Towards a neuroanatomy of autism: A systematic review and meta-analysis of structural magnetic resonance imaging studies. *Eur. Psychiatry* 2008; **23**: 289–299.
57. van Haren NEM, Schnack HG, Cahn W *et al.* Changes in cortical thickness during the course of illness in schizophrenia. *Arch. Gen. Psychiatry* 2011; **68**: 871–880.
58. Papmeyer M, Giles S, Sussmann JE *et al.* Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder. *Biol. Psychiatry* 2015; **78**: 58–66.
59. Ecker C, Ginestet C, Feng Y *et al.* Brain surface anatomy in adults with autism: The relationship between surface area, cortical thickness, and autistic symptoms. *JAMA Psychiat.* 2013; **70**: 59–70.
60. Hanford LC, Nazarov A, Hall GB, Sassi RB. Cortical thickness in bipolar disorder: A systematic review. *Bipolar Disord.* 2016; **18**: 4–18.
61. McLaughlin KA, Sheridan MA, Winter W, Fox NA, Zeanah CH, Nelson CA. Widespread reductions in cortical thickness following severe early-life deprivation: A neurodevelopmental pathway to attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 2014; **76**: 629–638.
62. Arat HE, Chouinard V-A, Cohen BM, Lewandowski KE, Ongür D. Diffusion tensor imaging in first degree relatives of schizophrenia and bipolar disorder patients. *Schizophr. Res.* 2015; **161**: 329–339.
63. Bracht T, Linden D, Keedwell P. A review of white matter microstructure alterations of pathways of the reward circuit in depression. *J. Affect. Disord.* 2015; **187**: 45–53.
64. Nortje G, Stein DJ, Radua J, Mataix-Cols D, Horn N. Systematic review and voxel-based meta-analysis of diffusion tensor imaging studies in bipolar disorder. *J. Affect. Disord.* 2013; **150**: 192–200.
65. Ameis SH, Catani M. Altered white matter connectivity as a neural substrate for social impairment in autism spectrum disorder. *Cortex* 2015; **62**: 158–181.
66. van Ewijk H, Heslenfeld DJ, Zwiers MP, Buitelaar JK, Oosterlaan J. Diffusion tensor imaging in attention deficit/hyperactivity disorder: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 2012; **36**: 1093–1106.
67. Schwerk A, Alves FDS, Pouwels PJW, van Amelsvoort T. Metabolic alterations associated with schizophrenia: A critical evaluation of proton magnetic resonance spectroscopy studies. *J. Neurochem.* 2014; **128**: 1–87.
68. Chitty KM, Lagopoulos J, Lee RSC, Hickie IB, Hermens DF. A systematic review and meta-analysis of proton magnetic resonance spectroscopy and mismatch negativity in bipolar disorder. *Eur. Neuropsychopharmacol.* 2013; **23**: 1348–1363.
69. Baruth JM, Wall CA, Patterson MC, Port JD. Proton magnetic resonance spectroscopy as a probe into the pathophysiology of autism spectrum disorders (ASD): A review. *Autism Res.* 2013; **6**: 119–133.
70. Selvaraj S, Arnone D, Cappai A, Howes O. Alterations in the serotonin system in schizophrenia: A systematic review and meta-analysis of postmortem and molecular imaging studies. *Neurosci. Biobehav. Rev.* 2014; **45**: 233–245.
71. Li Z, He Y, Tang J, Zong X, Hu M, Chen X. Molecular imaging of striatal dopamine transporters in major depression--a meta-analysis. *J. Affect. Disord.* 2015; **174**: 137–143.
72. Watkins CC, Sawa A, Pomper MG. Glia and immune cell signaling in bipolar disorder: Insights from neuropharmacology and molecular imaging to clinical application. *Transl. Psychiatry* 2014; **4**: e350.
73. Zürcher NR, Bhanot A, McDougle CJ, Hooker JM. A systematic review of molecular imaging (PET and SPECT) in autism spectrum disorder: Current state and future research opportunities. *Neurosci. Biobehav. Rev.* 2015; **52**: 56–73.
74. Swanson JM, Kinsbourne M, Nigg J *et al.* Etiologic subtypes of attention-deficit/hyperactivity disorder: Brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychol. Rev.* 2007; **17**: 39–59.
75. Ehlis A-C, Schneider S, Dresler T, Fallgatter AJ. Application of functional near-infrared spectroscopy in psychiatry. *Neuroimage* 2014; **85** (Pt 1): 478–488.
76. McLoughlin G, Makeig S, Tsuang MT. In search of biomarkers in psychiatry: EEG-based measures of brain function. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 2014; **165B**: 111–121.
77. Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neurosci. Biobehav. Rev.* 2012; **36**: 1140–1152.
78. Wolfers T, Buitelaar JK, Beckmann CF, Franke B, Marquand AF. From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neurosci. Biobehav. Rev.* 2015; **57**: 328–349.
79. Mamah D, Barch DM, Repovš G. Resting state functional connectivity of five neural networks in bipolar disorder and schizophrenia. *J. Affect. Disord.* 2013; **150**: 601–609.
80. Rashid B, Arbabshirani MR, Damaraju E *et al.* Classification of schizophrenia and bipolar patients using static and dynamic resting-state fMRI brain connectivity. *Neuroimage* 2016; **134**: 645–657.
81. Guerra-Carrillo B, Mackey AP, Bunge SA. Resting-state fMRI: A window into human brain plasticity. *Neuroscientist* 2014; **20**: 522–533.
82. Power JD, Schlaggar BL, Petersen SE. Studying brain organization via spontaneous fMRI signal. *Neuron* 2014; **84**: 681–696.
83. Shen H, Wang L, Liu Y, Hu D. Discriminative analysis of resting-state functional connectivity patterns of

schizophrenia using low dimensional embedding of fMRI. *Neuroimage* 2010; **49**: 3110–3121.
84. Fan Y, Liu Y, Wu H *et al.* Discriminant analysis of functional connectivity patterns on Grassmann manifold. *Neuroimage* 2011; **56**: 2058–2067.
85. Bassett DS, Nelson BG, Mueller BA, Camchong J, Lim KO. Altered resting state complexity in schizophrenia. *Neuroimage* 2012; **59**: 21196–21207.
86. Tang Y, Wang L, Cao F, Tan L. Identify schizophrenia using resting-state functional connectivity: An exploratory research and analysis. *Biomed. Eng. Online* 2012; **11**: 50.
87. Venkataraman A, Whitford TJ, Westin C-F, Golland P, Kubicki M. Whole brain resting state functional connectivity abnormalities in schizophrenia. *Schizophr. Res.* 2012; **139**: 7–12.
88. Anderson A, Cohen MS. Decreased small-world functional network connectivity and clustering across resting state networks in schizophrenia: An fMRI classification tutorial. *Front. Hum. Neurosci.* 2013; **7**: 520.
89. Arbabshirani MR, Kiehl KA, Pearlson GD, Calhoun VD. Classification of schizophrenia patients based on resting-state functional network connectivity. *Front. Neurosci.* 2013; **7**: 133.
90. Fekete T, Wilf M, Rubin D, Edelman S, Malach R, Mujica-Parodi LR. Combining classification with fMRI-derived complex network measures for potential neurodiagnostics. *PLoS One* 2013; **8**: e62867.
91. Su L, Wang L, Shen H, Feng G, Hu D. Discriminative analysis of non-linear brain connectivity in schizophrenia: An fMRI study. *Front. Hum. Neurosci.* 2013; **7**: 702.
92. Yu Y, Shen H, Zeng L-L, Ma Q, Hu D. Convergent and divergent functional connectivity patterns in schizophrenia and depression. *PLoS One* 2013; **8**: e68250–e68251.
93. Yu Y, Shen H, Zhang H, Zeng L-L, Xue Z, Hu D. Functional connectivity-based signatures of schizophrenia revealed by multiclass pattern analysis of resting-state fMRI from schizophrenic patients and their healthy siblings. *Biomed. Eng. Online* 2013; **12**: 10.
94. Watanabe T, Kessler D, Scott C, Angstadt M, Sripada C. Disease prediction based on functional connectomes using a scalable and spatially-informed support vector machine. *Neuroimage* 2014; **96**: 183–202.
95. Cheng H, Newman S, Goñi J *et al.* Nodal centrality of functional network in the differentiation of schizophrenia. *Schizophr. Res.* 2015; **168**: 345–352.
96. Chyzhyk D, Savio A, Graña M. Computer aided diagnosis of schizophrenia on resting state fMRI data by ensembles of ELM. *Neural Netw.* 2015; **68**: 23–33.
97. Kaufmann T, Skåtun KC, Alnæs D *et al.* Disintegration of sensorimotor brain networks in schizophrenia. *Schizophr. Bull.* 2015; **41**: 1326–1335.
98. Kim J, Calhoun VD, Shim E, Lee J-H. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *Neuroimage* 2016; **124**: 127–146.
99. Lord A, Horn D, Breakspear M, Walter M. Changes in community structure of resting state functional connectivity in unipolar depression. *PLoS One* 2012; **7**: e41282.
100. Wei M, Qin J, Yan R, Li H, Yao Z, Lu Q. Identifying major depressive disorder using Hurst exponent of resting-state brain networks. *Psychiatry Res.* 2013; **214**: 306–312.
101. Cao L, Guo S, Xue Z *et al.* Aberrant functional connectivity for diagnosis of major depressive disorder: A discriminant analysis. *Psychiatry Clin. Neurosci.* 2013; **68**: 110–119.
102. Guo H, Cheng C, Cao X, Xiang J, Chen J, Zhang K. Resting-state functional connectivity abnormalities in first-onset unmedicated depression. *Neural Regen. Res.* 2014; **9**: 153–163.
103. Fair DA, Nigg JT, Iyer S *et al.* Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. *Front. Syst. Neurosci.* 2012; **6**: 80.
104. Sidhu GS, Asgarian N, Greiner R, Brown MRG. Kernel principal component analysis for dimensionality reduction in fMRI-based diagnosis of ADHD. *Front. Syst. Neurosci.* 2012; **6**: 74.
105. Dey S, Rao AR, Shah M. Attributed graph distance measure for automatic detection of attention deficit hyperactive disordered subjects. *Front. Neural Circuits* 2014; **8**: 64.
106. Deshpande G, Wang P, Rangaprakash D, Wilamowski B. Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data. *IEEE Trans. Cybern.* 2015; **45**: 2668–2679.
107. Anderson JS, Nielsen JA, Froehlich AL *et al.* Functional connectivity magnetic resonance imaging classification of autism. *Brain* 2011; **134**: 3742–3754.
108. Nielsen JA, Zielinski BA, Fletcher PT *et al.* Multisite functional connectivity MRI classification of autism: ABIDE results. *Front. Hum. Neurosci.* 2013; **7**: 599.
109. Uddin LQ, Supekar K, Lynch CJ *et al.* Salience network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiat.* 2013; **70**: 869–879.
110. Chen CP, Keown CL, Jahedi A *et al.* Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism. *Neuroimage Clin.* 2015; **8**: 238–245.
111. Iidaka T. Resting state functional magnetic resonance imaging and neural network classified autism and control. *Cortex* 2015; **63**: 55–67.
112. Plitt M, Barnes KA, Martin A. Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *Neuroimage Clin.* 2015; **7**: 359–366.
113. Chen H, Duan X, Liu F *et al.* Multivariate classification of autism spectrum disorder using frequency-specific

resting-state functional connectivity—A multi-center study. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 2016; **64**: 1–9.
114. Yahata N, Morimoto J, Hashimoto R *et al.* A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nat. Commun.* 2016; **7**: 11254.
115. Di Martino A, Yan C-G, Li Q *et al.* The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 2014; **19**: 659–667.
116. Brown MRG, Sidhu GS, Greiner R *et al.* ADHD-200 global competition: Diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements. *Front. Syst. Neurosci.* 2012; **6**: 69.
117. Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 2014; **90**: 449–468.
118. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 2012; **59**: 2142–2154.
119. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* 2014; **84**: 320–341.
120. Girolami M. *Advances in Independent Component Analysis*. Springer-Verlag, London, Berlin, Heidelberg, 2000.
121. Fox MD, Zhang D, Snyder AZ, Raichle ME. The global signal and observed anticorrelated resting state brain networks. *J. Neurophysiol.* 2009; **101**: 3270–3283.
122. Chai XJ, Castañón AN, Öngür D, Whitfield-Gabrieli S. Anticorrelations in resting state networks without global signal regression. *Neuroimage* 2012; **59**: 1420–1428.
123. Whelan R, Garavan H. When optimism hurts: Inflated predictions in psychiatric neuroimaging. *Biol. Psychiatry* 2014; **75**: 746–748.
124. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. Springer Science & Business Media, New York, 2013.
125. Van Essen DC, Ugurbil K, Auerbach E *et al.* The human connectome project: A data acquisition perspective. *Neuroimage* 2012; **62**: 2222–2231.
126. Kennedy DN, Haselgrove C, Riehl J, Preuss N, Buccigrossi R. The NITRC image repository. *Neuroimage* 2016; **124**: 1069–1673.
127. Landis D, Courtney W, Dieringer C *et al.* COINS data exchange: An open platform for compiling, curating, and disseminating neuroimaging data. *Neuroimage* 2016; **124**: 1084–1088.
128. Mwangi B, Tian TS, Soares JC. A review of feature reduction techniques in neuroimaging. *Neuroinformatics* 2014; **12**: 229–244.
129. Krzanowski W. *Principles of Multivariate Analysis*. Oxford University Press, Oxford, 2000.
130. Dosenbach NUF, Nardos B, Cohen AL *et al.* Prediction of individual brain maturity using fMRI. *Science* 2010; **329**: 1358–1361.
131. Tomasi D, Volkow ND. Gender differences in brain functional connectivity density. *Hum. Brain Mapp.* 2011; **33**: 849–860.
132. Klaassens BL, van Gorsel HC, Khalili-Mahani N *et al.* Single-dose serotonergic stimulation shows widespread effects on functional brain connectivity. *Neuroimage* 2015; **122**: 440–450.
133. Friedman L, Glover GH. The FBIRN consortium. Reducing interscanner variability of activation in a multicenter fMRI study: Controlling for signal-to-fluctuation-noise-ratio (SFNR) differences. *Neuroimage* 2006; **33**: 471–481.
134. Thulborn KR, Chang SY, Shen GX, Voyvodic JT. High-resolution echo-planar fMRI of human visual cortex at 3.0 tesla. *NMR Biomed.* 1997; **10**: 183–190.
135. Van Dijk KRA, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *J. Neurophysiol.* 2010; **103**: 297–321.
136. Mennes M, Biswal BB, Castellanos FX, Milham MP. Making data sharing work: The FCP/INDI experience. *Neuroimage* 2013; **82**: 683–691.
137. Keator DB, van Erp TGM, Turner JA *et al.* The function biomedical informatics research network data repository. *Neuroimage* 2016; **124**: 1074–1079.
138. Witten DM, Tibshirani R, Hastie T. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics* 2009; **10**: 515–534.
139. Yamashita O, Sato M-A, Yoshioka T, Tong F, Kamitani Y. Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage* 2008; **42**: 1414–1429.
140. Nieuwenhuis M, van Haren NEM, Pol HEH, Cahn W, Kahn RS, Schnack HG. Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *Neuroimage* 2012; **61**: 606–612.
141. Schnack HG, Nieuwenhuis M, van Haren NEM *et al.* Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *Neuroimage* 2014; **84**: 299–306.
142. Pinkham AE, Hopfinger JB, Pelphrey KA, Piven J, Penn DL. Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophr. Res.* 2008; **99**: 164–175.
143. King BH, Lord C. Is schizophrenia on the autism spectrum? *Brain Res.* 2011; **1380**: 34–41.
144. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.

145. Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. *Science* 2015; **349**: 1489–1494.
146. Insel TR, Cuthbert BN. Brain disorders? Precisely. *Science* 2015; **348**: 499–500.
147. Connor JP, Symons M, Feeney GFX, Young RM, Wiles J. The application of machine learning techniques as an adjunct to clinical decision making in alcohol dependence treatment. *Subst. Use Misuse* 2007; **42**: 2193–2206.
148. Nejad AB, Madsen KH, Ebdrup BH *et al.* Neural markers of negative symptom outcomes in distributed working memory brain activity of antipsychotic-naive schizophrenia patients. *Int. J. Neuropsychopharmacol.* 2013; **16**: 1195–1204.
149. Perlis RH. A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. *Biol. Psychiatry* 2013; **74**: 7–14.
150. Amminger GP, Mechelli A, Rice S *et al.* Predictors of treatment response in young people at ultra-high risk for psychosis who received long-chain omega-3 fatty acids. *Transl. Psychiatry* 2015; **5**: e495.
151. Kim J-W, Sharma V, Ryan ND. Predicting methylphenidate response in ADHD using machine learning approaches. *Int. J. Neuropsychopharm.* 2015; **18**: pyv052.
152. Månsson KNT, Frick A, Boraxbekk C-J *et al.* Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Transl. Psychiatry* 2015; **5**: e530.
153. Patel MJ, Andreescu C, Price JC, Edelman KL, Reynolds CF, Aizenstein HJ. Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. *Int. J. Geriatr. Psychiatry* 2015; **30**: 1056–1067.
154. Yun J-Y, Jang JH, Kim SN, Jung WH, Kwon JS. Neural correlates of response to pharmacotherapy in obsessive–compulsive disorder: Individualized cortical morphology-based structural covariance. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 2015; **63**: 126–133.
155. Iniesta R, Malki K, Maier W *et al.* Combining clinical variables to optimize prediction of antidepressant treatment outcomes. *J. Psychiatr. Res.* 2016; **78**: 94–102.
156. Koutsouleris N, Meisenzahl EM, Davatzikos C *et al.* Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Arch. Gen. Psychiatry* 2009; **66**: 700–712.
157. Galatzer-Levy IR, Karstoft K-I, Statnikov A, Shalev AY. Quantitative forecasting of PTSD from early trauma responses: A machine learning application. *J. Psychiatr. Res.* 2014; **59**: 68–76.
158. Karstoft K-I, Statnikov A, Andersen SB, Madsen T, Galatzer-Levy IR. Early identification of posttraumatic stress following military deployment: Application of machine learning methods to a prospective study of Danish soldiers. *J. Affect. Disord.* 2015; **184**: 170–175.
159. Schmaal L, Marquand AF, Rhebergen D *et al.* Predicting the naturalistic course of major depressive disorder using clinical and multimodal neuroimaging information: A multivariate pattern recognition study. *Biol. Psychiatry* 2015; **78**: 278–286.
160. Seo S, Mohr J, Beck A, Wüstenberg T, Heinz A, Obermayer K. Predicting the future relapse of alcohol-dependent patients from structural and functional brain images. *Addict. Biol.* 2015; **20**: 1042–1055.
161. Rush AJ, Fava M, Wisniewski SR *et al.* Sequenced treatment alternatives to relieve depression (STAR*D): Rationale and design. *Control. Clin. Trials* 2004; **25**: 119–142.
162. Thibault RT, Lifshitz M, Raz A. The self-regulating brain and neurofeedback: Experimental science and clinical promise. *Cortex* 2016; **74**: 247–261.
163. Thibault RT, Lifshitz M, Birbaumer N, Raz A. Neurofeedback, self-regulation, and brain imaging: Clinical science and fad in the service of mental disorders. *Psychother. Psychosom.* 2015; **84**: 193–207.
164. Berman BD, Horovitz SG, Hallett M. Modulation of functionally localized right insular cortex activity using real-time fMRI-based neurofeedback. *Front. Hum. Neurosci.* 2013; **7**: 638.
165. Emmert K, Breimhorst M, Bauermann T, Birklein F, Van De Ville D, Haller S. Comparison of anterior cingulate vs. insular cortex as targets for real-time fMRI regulation during pain stimulation. *Front. Behav. Neurosci.* 2014; **8**: 350.
166. Brühl AB, Scherpiet S, Sulzer J, Stämpfli P, Seifritz E, Herwig U. Real-time neurofeedback using functional MRI could improve down-regulation of amygdala activity during emotional stimulation: A proof-of-concept study. *Brain Topogr.* 2013; **27**: 138–148.
167. Paret C, Kluetsch R, Ruf M *et al.* Down-regulation of amygdala activation with real-time fMRI neurofeedback in a healthy female sample. *Front. Behav. Neurosci.* 2014; **8**: 299.
168. Johnston S, Linden DEJ, Healy D, Goebel R, Habes I, Boehm SG. Upregulation of emotion areas through neurofeedback with a focus on positive mood. *Cogn. Affect. Behav. Neurosci.* 2010; **11**: 44–51.
169. Megumi F, Yamashita A, Kawato M, Imamizu H. Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Front. Hum. Neurosci.* 2015; **9**: 160.
170. Lai C-Y, Scarr E, Udawela M, Everall I, Chen WJ, Dean B. Biomarkers in schizophrenia: A focus on blood based diagnostics and theranostics. *World J. Psychiatry* 2016; **6**: 102–117.
171. Goel S, England CG, Chen F, Cai W. Positron emission tomography and nanotechnology: A dynamic duo for cancer theranostics. *Adv. Drug Deliv. Rev.* 2016. doi: 10.1016/j.addr.2016.08.001