Clinical EEG and Neuroscience
1–14

DOI: 10.1177/1550059418782093
journals.sagepub.com/home/eeg
SAGE

Multimodal Neuroimaging: Basic Concepts and Classification of Neuropsychiatric Diseases

Emine Elif Tulay[1], Barış Metin[1], Nevzat Tarhan[1,2], and Mehmet Kemal Arıkan[1]

Abstract

Neuroimaging techniques are widely used in neuroscience to visualize neural activity, to improve our understanding of brain mechanisms, and to identify biomarkers—especially for psychiatric diseases; however, each neuroimaging technique has several limitations. These limitations led to the development of multimodal neuroimaging (MN), which combines data obtained from multiple neuroimaging techniques, such as electroencephalography, functional magnetic resonance imaging, and yields more detailed information about brain dynamics. There are several types of MN, including visual inspection, data integration, and data fusion. This literature review aimed to provide a brief summary and basic information about MN techniques (data fusion approaches in particular) and classification approaches. Data fusion approaches are generally categorized as asymmetric and symmetric. The present review focused exclusively on studies based on symmetric data fusion methods (data-driven methods), such as independent component analysis and principal component analysis. Machine learning techniques have recently been introduced for use in identifying diseases and biomarkers of disease. The machine learning technique most widely used by neuroscientists is classification—especially support vector machine classification. Several studies differentiated patients with psychiatric diseases and healthy controls with using combined datasets. The common conclusion among these studies is that the prediction of diseases increases when combining data via MN techniques; however, there remain a few challenges associated with MN, such as sample size. Perhaps in the future N-way fusion can be used to combine multiple neuroimaging techniques or nonimaging predictors (eg, cognitive ability) to overcome the limitations of MN.

Keywords

multimodal neuroimaging, fusion, machine learning, classification, psychiatry

Received February 16, 2018; revised May 15, 2018; accepted May 17, 2018.

Introduction

Especially in the past 20 years, neuroimaging made quite a splash because of the improvements in computing technology and improved our understanding of the mechanisms of brain as well as our ability to detect the cause of impairment via classification of patients and healthy controls. Furthermore, neuroimaging techniques are invaluable for identifying potential neurobiological markers and generating predictions for preventing the progression of various diseases.

Generally, neuroimaging is an umbrella term for multiple methods, technologies, and noninvasive techniques (modalities) that provide structural and functional data regarding neural mechanisms, including electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and near-infrared spectroscopy (NIRS), which yield functional information (e.g. neural activity and cognitive functions), computed tomography (CT), structural MRI (sMRI), and diffusion tensor imaging (DTI), which yield structural/anatomical information (eg, gray matter and white matter tracts). Each of these techniques has advantages and disadvantages related to resolution, safety, availability, and accessibility.[1] For example, EEG has high temporal, but low spatial resolution, whereas fMRI has high spatial, but low temporal resolution (see next section). Moreover, there are additional techniques used with neuroimaging techniques for giving stimulation such as transcranial magnetic stimulation (TMS) and for source localization problem of EEG such as variable resolution electromagnetic tomography (VARETA), low-resolution brain electromagnetic tomography

[1]Uskudar University, Istanbul, Turkey
[2]NPIstanbul Hospital, Istanbul, Turkey

Corresponding Author:
Mehmet Kemal Arıkan, Department of Psychology, Faculty of Humanities and Social Sciences, Uskudar University, Istanbul, Turkey.
Email: mkarikan46@gmail.com

(LORETA) with different methods such as standardized LORETA (sLORETA) and exact LORETA (eLORETA).

Although each modality provides different and valuable information about brain structures and/or activity, researchers began combining multiple techniques, referred to as multimodal neuroimaging (MN), to compensate for the limitations of each modality, so as to understand brain dynamics with greater detail (see next section). Generally, there are 3 approaches to MN: (*a*) visual inspection, (*b*) data integration, and (*c*) data fusion[2] (see section "Principles of Data Fusion"). Nonetheless, according to Correa et al,[3] unlike data integration methods, data fusion facilitates true interaction between different types of data. In literature, data fusion can be categorized as asymmetric or symmetric, and each category uses a variety of techniques, including principal component analysis (PCA), independent component analysis (ICA), and general linear models (see section "Principles of Data Fusion").

Interpretation of findings and identification of biomarkers, especially for neuropsychiatric diseases, is not always an easy process, regardless of the use of unimodal neuroimaging (single neuroimaging technique) or MN[4]; therefore, neuroimaging studies that use machine learning (ML) as a prognostic/diagnostic tool are becoming more common. Classification is one of the ML techniques used for modeling (decoding) and predicting categorical variables and includes different methods such as support vector machine (SVM) classification, which is the most commonly used method among other classification methods. The other MN techniques are regression and clustering (see section "Decoding Mental States Based on Classification").

This literature review aimed to (*a*) provide a brief summary and basic information about MN techniques (data fusion approaches in particular) and classification approaches, (*b*) describe various techniques of data fusion and classification for use in psychiatry, and (*c*) provide an overview of how data obtained from multiple imaging techniques (eg, EEG and fMRI) are combined via data fusion and how psychiatric diseases are classified using such combined data. The literature was searched via PubMed, Science Direct, Web of Science, and Google using the following keywords: multimodal; neuroimaging; fusion; data integration; univariate; multivariate; psychiatry; neuropsychiatric diseases; mental disorder; bipolar; schizophrenia; psychosis; attention-deficit hyperactivity disorder; major depressive disorder; depression; classification; machine learning; pattern recognition; accuracy. The search also focused on studies published after 2010 that differentiated patients and controls based on MN and data fusion.

The following section includes a brief description of unimodal neuroimaging (single-technique neuroimaging) techniques, an informative introduction to MN and different approaches for combining data obtained using multiple unimodal techniques, and a survey of studies based on MN techniques—especially psychiatry studies. The subsequent section provides an overview of classification techniques used for decoding brain activity and of studies that differentiated patients and healthy controls based on single neuroimaging modality. In addition, studies on the accuracy of MN are reviewed, so as to illustrate the effectiveness of the data fusion process.

Application of Neuroimaging: From Unimodal to Multimodal

The unimodal concept refers to use of a single neuroimaging technique that measures electrophysiological or hemodynamic signals. The literature includes many comprehensive reviews that explain these neuroimaging techniques in detail, including their advantages and limitations[1,5-14]; an overview of these techniques is presented in Table 1.

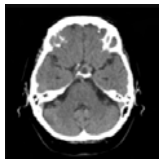
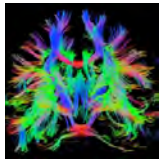
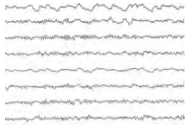
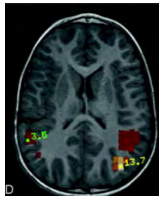
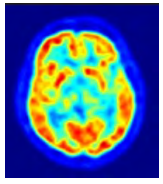
In addition, the literature includes many studies on neuropsychiatric diseases based on various unimodal techniques. A review by Phillips and Swartz[15] describes several studies on bipolar disorder (BD) that used fMRI, volumetric analysis, DTI, and resting state techniques. There is also study that compared EEG data in BD during manic and depressive episodes.[16] Furthermore, several studies focused on neuroimaging biomarkers for major depressive disorder (MDD)[17] and schizophrenia (SZ),[18,19] and comparing patients diagnosed with SZ and MDD,[20] using sLORETA. An informative overview of neuroimaging techniques for all neuropsychiatric diseases can be found in Malhi and Lagopoulos[6] and Hughes and John[21]; however, as mentioned earlier each neuroimaging modality has specific technological and physiological limitations that are leading to more widespread use of MN among neuroscientists. Bießmann[8] presents an extensive description of the progression of neuroimaging from unimodal neuroimaging to MN with a history flow that includes important advancements in neuroscience.

What Is Multimodal Neuroimaging?

In general, MN is an approach that combines data sets obtained using $\geqslant 2$ unimodal modalities, such as EEG and fMRI integration (referred to as neurovascular coupling),[8,22] which is the most common MN approach,[23-25] to yield more informative, consistent, and reliable results than can be obtained using unimodal neuroimaging. Uludağ and Roebroeck[26] define MN in 2 terms: (*a*) narrow sense and (*b*) wider sense. In the narrow sense, MN refers to the combination of data obtained from different instruments (separately recorded modalities). In this sense, the combination can be between modalities that either separately analyzed or jointly analyzed (see section "Decoding Mental States Based on Classification"). In the wider sense, MN is defined as the combination of data recorded with the same physical instrument (simultaneously recorded).

Although MN poses its own challenges, such as sample size and number of dimensions (see Lahat et al,[27] Bießmann et al,[28] and references therein), it has several advantages over unimodal neuroimaging, including higher spatial and temporal resolution, and provision of more comprehensive information regarding neural processes, structures, quantification, generalization, and normalization,[26,28] and overcomes the limitations of unimodal techniques.[8] Thus, MN can play an important role

Table 1. Advantages and Limitations of Neuroimaging Techniques.

Methods		Measurement Provided	Temporal Resolution	Spatial Resolution	Advantages	Limitations
	CT	Brain structures	Minutes	0.5-1 mm	High spatial resolution	Radiation Low contrast Low temporal resolution
	MRI	Brain structures (eg, white matter, gray matter, and cerebrospinal fluid)	Minutes to hours	1-2 mm	High spatial resolution No radiation	Low temporal resolution Relatively low sensitivity High cost Long scanning time
	DTI	Fiber tracks	Minutes	2.5 mm	High spatial resolution	Limited information for GM
	EEG	Brain activity	Milliseconds	>10 mm	High temporal resolution No radiation Low cost Portable Widely available Fewer motion artefacts	Low spatial resolution Does not measure activity below the cortex
	MEG	Brain magnetic activity	Milliseconds	>5 mm	High temporal resolution Medium spatial resolution	Low spatial resolution Not portable Limited availability High cost
	PET	Perfusion Metabolism Neurotransmitter dynamics	Seconds to minutes	4-10 mm	Fewer motion artifacts High sensitivity	Low spatial and temporal resolution Limited availability Radiation High cost Not portable

(continued)

Table 1. (continued)

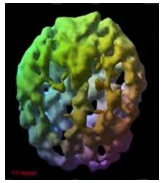
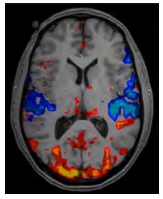
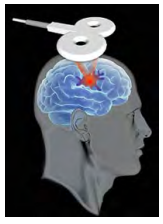
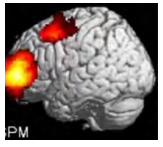
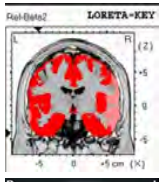
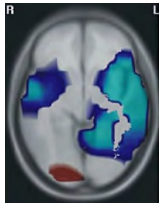
Methods		Measurement Provided	Temporal Resolution	Spatial Resolution	Advantages	Limitations
	SPECT	Perfusion Metabolism Neurotransmitter dynamics	Minutes	8-15 mm	High sensitivity Lower cost than PET Higher availability than PET	Low spatial and temporal resolution Lower sensitivity than PET
	fMRI	Hemodynamic activity	Seconds	< 3 mm	High spatial resolution No radiation Widely available	Not portable Low temporal resolution Sensitive to motion artefacts
	TMS[a]	Focal brain activity	Milliseconds to seconds	45-90 mm	No radiation Portable Can stimulate lesions	Spatial and temporal resolution dependent on other parameters Has some risks (eg, seizures, damage brain cells)
	NIRS	Fluctuations in cerebral metabolism during neural activity	Seconds	> 5 mm	Medium temporal resolution Low cost Portable	Low spatial resolution
 	LORETA/VARETA	Brain electric/magnetic activity	Milliseconds	5-7 mm	High estimation accuracy of the current density and location Low error rate High time resolution VARETA imposes different amounts of spatial smoothness for different types of generators. VARETA eliminate ghost solutions and minimize the diffuse allocation of variance	Low spatial resolution when compared with that of an fMRI or PET scan Need some algorithms for spatial blurring In LORETA, regularization parameter is a constant

Table 2. Advantages and Limitations/Challenges/Bias of the Multimodal Neuroimaging Approach.

Advantages	Limitations/Challenges/Bias
• Exploratory	• Noncommensurable
• Robust and redundant	• Different resolutions
• Unique and identifiable solution	• Number of dimensions
• High spatiotemporal resolution	• Inconsistent data
• Improved data quality	• Sample size

in the detection, diagnosis, prognosis, and treatment of some diseases like neuropsychiatric diseases[1] (Table 2).

With MN there are multiple ways to combine data obtained from different unimodal modalities. Calhoun and Sui[2] categorized MN approaches as follows: (*a*) visual inspection: unimodal analysis results are visualized separately; (*b*) data integration: data obtained with each unimodal technique are analyzed individually and then overlaid, which prevents any interaction between different types of data[29]; (*c*) data fusion: one modality constrains another modality (asymmetric data fusion) or all modalities are analyzed jointly (symmetric data fusion) (for more details about data fusion, see section "Decoding Mental States Based on Classification"). The most qualitative data are obtained via the data fusion approach, followed by data integration and visual inspection.[2,7]

Liu et al[1] documented the rapid progression of MN research from 1975 to 2014 in an extensive literature review. Published studies have used various combinations of 2 or more unimodal modalities, such as structural-structural (eg, sMRI + DTI), functional-functional (eg, EEG + fMRI), and structural-functional (eg, fMRI + DTI); related studies can be found in Sui et al,[30,31] Bießmann et al,[28] Calhoun and Sui,[2] and Schultz et al,[32] and Ahn and Jun.[33]

Principles Of data Fusion

Lahat et al[27] define data fusion as, "the analysis of several datasets, such that different datasets can interact and inform each other." According to Calhoun and Adalı,[7] data fusion is a process that utilizes multiple image types simultaneously in order to take advantage of the cross-information. Simply stated, data fusion is the analysis of $\geqslant$2 brain imaging modalities collectively.[3]

Calhoun and Sui[2] show cumulatively increment for usage of data fusion, including 2-way and N-way fusion, which refers to a combination of $\geqslant$2 modalities, where N is the number of modalities.[34] Furthermore, Wolfers et al[35] emphasize the importance of combining data from multiple sources in cases of psychiatric diseases that are affected by multiple factors. In contrast, Lahat et al[36] highlighted some of the challenges associated with data fusion, including data-related problems (eg, different resolution, inconsistent data), level of data fusion (eg, data integration), model, and theoretical validation.

Data fusion methods are generally divided into 2 groups: symmetric and asymmetric. With the asymmetric data fusion approach data obtained using 1 modality are used to guide the analysis of data obtain via another modality.[37] For example, data obtained via EEG can be used as a regressor in the analysis of fMRI data in order to extract voxels that correlate with the EEG regressors (EEG constrains fMRI analysis), or fMRI constrains EEG source localization problem with the spatial information (fMRI constrains EEG analysis) (see He and Liu[38] and references therein). On the other hand, the most commonly used data fusion approach for MN is symmetric data fusion, which is used to simultaneously analyze data sets collected using multiple modalities. Symmetric data fusion is subgrouped (Figure 1) as modal driven (hypothesis driven)[39] and data driven, both of which include a variety of fusion methods (for more information, see Valdes-Sosa et al[39]).

Data fusion methods can be univariate (eg, correlation and *t* tests) or multivariate (eg, ICA and PCA). Univariate pattern analysis is used to examine the mean difference between 2 conditions, so as to understand whether or not there is consistency across patients[40] whereas multivariate pattern analysis (MVPA) is used to identify correlated patterns (components) between multiple datasets (obtained via $\geqslant$2 modalities). MVPA (also known as multivoxel pattern analysis, in the context of fMRI analysis) has some advantages over the univariate approach; for example, it provides robustness to noise.[2]

Although individual predictions can be made using univariate techniques (SZ,[41] BD[42,43]), MVPA can integrate various data in an efficient way, in addition to identify biomarkers.[35] Various methods are used for MVPA, including PCA,[44] joint ICA (jICA),[2,45] parallel ICA (pICA),[46,47] canonical correlation analysis (CCA),[3] temporal kernel canonical correlation analysis (tkCCA),[48] partial least squares (PLS),[49] linked ICA,[50] LASSO (least absolute shrinkage and selection operator),[51] and coefficient-constrained ICA (CC-ICA),[52] as well as combinations of these methods, including mCCA + jICA (SZ[29,53,54], obsessive-compulsive disorder [OCD][55]). As this literature review did not aim to explain these methods, more details can be found in the cited studies. Additionally, the literature includes several extensive reviews that mention listed multivariate analyses above and other more.[51,56]

Decoding Mental States Based on Classification

Decoding aids predicting the course of diseases using brain signals.[57,58] For this purpose, a model is used to examine significant differences, for example between patients and healthy controls. This model can be based on simple statistical methods (eg, grand averages and between-group differences)[59] or more complicated ML algorithms (eg, regression analysis and classification algorithms).[60] Although some challenges (such as sample size) remain,[60] interest in the use of ML algorithms for decoding brain activity continues to increase.[61,62]

ML is a common name for several algorithms which identify patterns in data for making predictions.[63] These algorithms are generally grouped into 2 methodological categories: supervised and unsupervised (Figure 2).[64] Supervised learning algorithms use known (predefined) input and output data, and then

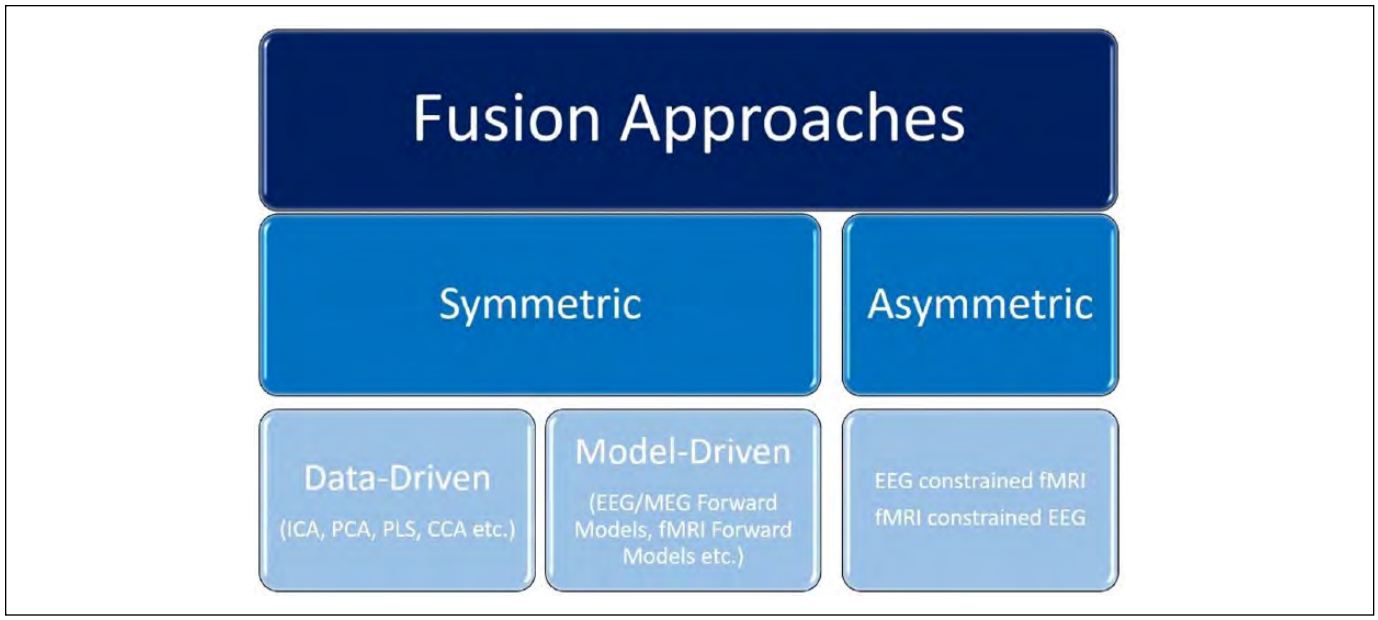


Figure 1. Data fusion techniques.
Abbreviations: EEG, electro encephalography; fMRI, functional magnetic resonance imaging; MEG, magnetoencephalography; ICA, independent component analysis; PCA, principal component analysis; CCA, canonical correlation analysis; PLS, partial least squares.

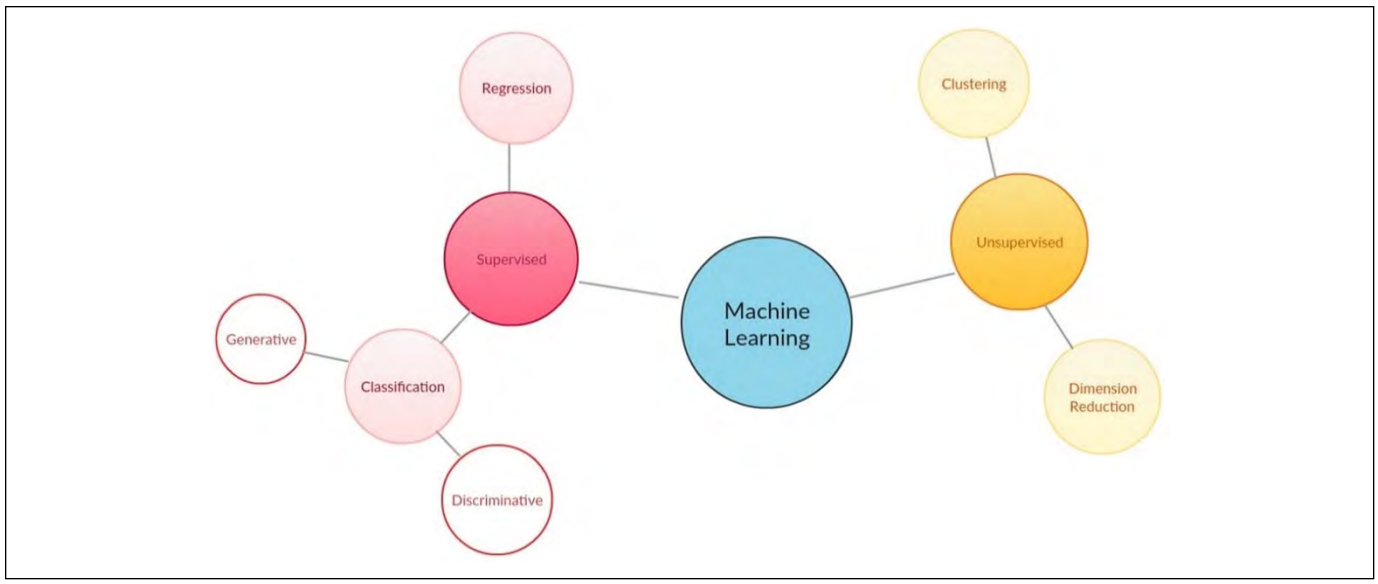


Figure 2. Categorization of machine learning techniques.

train a model to generate reasonable predictions about the response to new data. Conversely, unsupervised learning algorithms does not know what the data (not predefined) for attempting to identify patterns and are most commonly used to identify hidden patterns in data.

According to the literature, supervised learning classification techniques are primarily used for making predictions (eg, different diseases) rather than regression analysis (eg, general linear models, decision trees, PLS, and linear regression). Various classifiers can be used to make a prognosis or diagnosis, the most commonly used classifiers are as follows: generative models—SVM,[65] deep learning,[66] and logistic regression[67]; discriminative models—Gaussian process classifiers,[68] multiple stepwise discriminant analyses,[69] and linear discriminant analysis.[70] There is a comprehensive review[71] that mentioned advantages and limitations of various classification techniques in using bioinformatics and neuroimaging.

There is a lack of consensus concerning how to choose the most appropriate classifier,[64] hence there are several studies have attempted to determine which is the most powerful

classification method for identifying the classes by comparing the performance of multiple classifiers.[72-75]

The Basic Classification Process

The classification process includes several steps: feature extraction, feature selection (reduction), and classification (training and testing/validation).[35,63-65,76-78] Feature extraction involves the transformation of original data into a form (feature* vector) that is meaningful to the classifier; this step is mandatory. During the feature selection (reduction) step, more important and/or redundant features are selected for differentiation between classes. Although this step has the potential to improve classification performance,[64,79] it is optional. Feature selection can be accomplished using various methods (see section "Principles of Data Fusion"), but the most commonly used methods are PCA and ICA, both of which are dimensionality reduction techniques. PCA extracts the most important characteristics from data and ICA identifies the components of the data that are mutually independent. Classification consists of 2 substeps: training and testing (validation). Training is used to teach an application to correctly classify data. All classifiers listed above can be used, but SVM[62,65] is widely considered the most powerful training method. For the testing substep, cross-validation methods (eg, leave-one-out cross-validation/jackknife, k-fold cross validation, and holdout)[64] are used to estimate how well a model has been trained. Sensitivity, specificity, and accuracy are the most commonly reported measures[63] a of classifier's performance,[†] and sample size is a very important parameter related to measurement accuracy; as sample size increases measurement accuracy decreases.[80]

Multimodal Neuroimaging and Classification: Classification of Psychiatric Diseases Based on MN and Data Fusion Approaches

In recent years, the number of studies based on supervised learning algorithms, especially in a classification framework (eg, SVM, Gaussian naive Bayes, and artificial neural network), for the prognosis and diagnosis of diseases has been increasing (neuropsychiatric diseases—mixed[35,60,65,82,83]; depression[84]; MDD[85-87]; uni/bipolar depressive disorder[88]; anxiety disorder[89]; social anxiety disorder (SAD)[90]; ADHD[63,91-98]; and references therein). There are also several studies using regression algorithms (healthy controls[99]; childhood autism[100]; MDD[101]; SZ[102]). Unsupervised algorithms consist of clustering algorithms (eg, K-means cluster)[103-105] and dimension reduction algorithms (eg, ICA, PCA) (related studies are provided in section "Principles of Data Fusion").

Although numerous studies in the literature used a single modality for the classification of several diseases, the present review focused on studies that used MN for classification of psychiatric disorders (Table 3). There have been several studies that aimed to differentiate SZ patients from healthy controls (HC) by combining data from rs-fMRI/task-related fMRI, and sMRI,[106-109] fMRI and single nucleotide polymorphism (SNP; genetic factor),[110,111] and rs-fMRI and MEG,[112] and some studies combined data from different 3 modalities[34,113] (accuracy ranged from 75% to 100%).

Ford et al[109] classified SZ and HC via Fisher's linear discriminate classifier by using task-related fMRI activation with 78% accuracy and sMRI data with 52% accuracy but the best accuracy (87%) was obtained by using combined data (activation + volume). However, the study had a small sample size for classification and testing (validation). Yang et al[108] combined connectivity features from rs-fMRI and anatomical features of sMRI data selected by ICA. They, then applied SVM for classification. Their findings show that combination of modalities (77.91%) yielded higher accuracy than using a single modality (72.09%). Cabral et al[107] classified SZ patients based on sMRI data with 69.7% accuracy, versus accuracy of 70.5% based on rs-fMRI data, and 75% accuracy was obtained when sMRI and rs-fMRI data were combined. Qureshi et al[106] developed further on this former study using the combination of rs-fMRI and sMRI data and increasing sample size. Using ELM classifiers, they obtained 99.29% accuracy Although the results of all studies are convincing for the use of combined data, the accuracy rates could have been better if the EEG or MEG methods that has higher temporal resolution had been used as an additional modality.

Another method used for classifying SZ and HC is combining fMRI and genetic data (eg, SNPs). Yang et al[111] used ICA and SVM-based classifier ensemble (SVME) methods for classification and measured accuracies for (*a*) SNP data alone (SNP-SVME), (*b*) fMRI activations alone(voxel-SVME), (*c*) components of fMRI activation obtained with ICA followed by SVM (ICA-SVMC), and (*d*) integration of fMRI and SNP data (combined SNP-fMRI). The accuracies they obtained were 73.88% for SNP-SVME, 81.63% for voxel-SVME, 82.50% for ICA-SVMC, and 87.25% for combined SNP-fMRI. However, their sample size and the size of SNP array (dataset of genotypes) were relatively small. On the other hand, Cao et al[110] analyzed a large dataset for distinguishing SZ from HC. They combine fMRI and SNP data with a model named as generalized sparse model (GSM) in which they selected the features by sparse representation-based variable selection (SRVS) algorithm with four models. They compared several classifiers, including sparse representation-based classifier (SRC), fuzzy c-means (FCM) classifier, and SVM-based classifier, and the best results obtained with SRC with 89.7% accuracy. Although combining the neuroimaging techniques with SNP seems to be

*Feature is a characteristic that is extracted from data; for example, voxels obtained from fMRI data.[35]

[†]The clear descriptions of the terms sensitivity, specificity, and accuracy were made by O'Halloran et al,[81] as follows, "In the case of binary classifiers, for example, involving patients and controls, sensitivity refers to the proportion of patients (true positives) who are correctly identified as patients, whereas specificity measures the proportion of controls (true negatives) who are correctly identified as controls. The accuracy of the classifier refers to the total proportion of patients and controls that are correctly classified."

Table 3. Overview of Studies on the Classification of Psychiatric Diseases Based on Multimodal Neuroimaging and Fusion Techniques.

Study	Participants	Modalities	Features	Methods	Accuracy
Schizophrenia					
Ford et al (2002)	8 HC 15 SZ	Task fMRI sMRI	HF voxels HF volume	PCA FLD	78% for fMRI 52% for sMRI 87% with combination of modalities
Yang et al (2010)	20 HC 20 SZ	Task fMRI SNP	Voxels in the fMRI map SNPs	ICA SVM-based classifier ensemble (SVME)	73.88% with SNP 81.63% with fMRI 87.25% with combination of modalities
Sui et al (2013)	45 HC 52 SZ	rs-fMRI sMRI DTI	ALFF GM density FA	mCCA + jICA LSVM RSVM KNN GNB	The most powerful prediction (>90% accuracy) can be accomplished using features from FA + GM via RSVM
Sui et al (2014)	53 HC 48 SZ	rs-fMRI sMRI EEG	ALFF of rs-fMRI GM segmentation image from sMRI EEG spectra	mCCA + jICA Combination of 2-sample t-test, MCCA, SVM-RFE	74% in training and 80% predication rate for EEG 84% in training and 90% predication rate for fMRI 86% in training and 80% predication rate sMRI 91% in training and 100% predication rate with combination of modalities
Cao et al (2014)	116 HC 92 SZ	fMRI SNP	Voxels SNPs	GSM SRVS	89.7% with combination of modalities
Yang et al (2016)	46 HC 40 SZ	rs-fMRI sMRI	FC Anatomical features of sMRI	ICA SVM	77.91% with combination of modalities
Cabral et al (2016)	74 HC 71 SZ	rs-fMRI sMRI	Connectivity features of fMRI Anatomical features (GM volume) of sMRI	PCA ν-SVR	69.7% with sMRI 70.5% with rs-fMRI 75% with combination of modalities
Çetin et al (2016)	44 HC 47 SZ	rs-fMRI MEG	FC MEG data for each frequency	Sg-ICA LDC NBC non-linear SVM	The average value of 3 classification methods' accuracies for dynamic functional network connectivity 82.79% for fMRI 67.03% for ensemble of MEG features 87.91% with combination of modalities
Qureshi et al (2017)	72 HC 72 SZ	rs-fMRI sMRI	FC Different features of sMRI	ICA ELM linear and non-linear (radial basis function), SVM, LDA, random forest ensemble	99.29% (ELM classifiers) with combination of modalities
Psychosis					
Pettersson-Yeo et al (2014)	23 HC 19 UHR 19 FEP	Task fMRI sMRI DTI	Different features of fMRI GM FAS	SK MKL AV MV SVM	86.33% with combination of DTI and fMRI for classifying FEP from UHR 83.33% with combination of all modalities for classifying FEP from UHR
Major Depressive Disorder and Depression					
Ota et al (2013)	Exploration sample 25 SZ 25 MDD validation sample 18 SZ 16MDD	sMRI DTI	GM Volume, ventricle volume FA	Stepwise discriminant analysis	72% for SZ 88% for MDD
Schmaal et al (2015)	23 chronic MDD 36 gradual-improving MDD 59 fast remission MDD	Task fMRI sMRI	Features of fMRI GM	Binary GPC	62% chronic vs. remitted 61% chronic vs. gradually improved 44% gradually improved vs. remitted
Schnyer et al (2017)	25 HC 25 MDD	sMRI DTI	WM FA	TBSS SVM	Whole-brain FA map total classification accuracy was 70.0% 74% for brain map of white matter fractional anisotropy values (FA)

(continued)

Table 3. (continued)

Study	Participants	Modalities	Features	Methods	Accuracy
ADHD					
Bohland et al (2012)	482 HC 272 ADHD	Rs-fMRI sMRI Phenotypic data	FC WM, CSF IQ-related phenotypic features	2-sample t -test Linear SVM	74% for sMRI For fMRI 67% CORR Network 71% SIC Network 61% KAPPA Network 76% for combination of sMRI and fMRI features
Colby et al (2012)	491 HC 285 ADHD	rs-fMRI sMRI	Different features of fMRI Different features of sMRI	SVM-RFE RBF-SVM	55% for combined data
Dai et al (2012)	402 HC 222 ADHD	rs-fMRI sMRI	ReHo, FC CT, GM	Combination of filter-based and wrapper-based methods SVM-RFE, MKL	61.54% by 2-class classifier for combined data
Anderson et al (2014)	472 HC 276 ADHD	rs-fMRI sMRI Phenotypic data	FC GM IQ-related phenotypic features	NMF, ICA Decision tree	66.8%
Qureshi et al (2017) (meta-analysis)	53 HC 53 ADHDI 53 ADHDC	rs-fMRI sMRI	FC Different features of sMRI	LASSO Different classifiers (ELM, ELM-NFS, SVM Linear, SVM-RBF) and binary classification	76.190% accuracy for ELM in multi-class settings, 73.81% accuracy for sMRI classification, 71.429% accuracy for fMRI classification 92.857% accuracy between ADHDI-HC based on binary classification

Abbreviations: HC, healthy controls; SZ, schizophrenia; rs-fMRI, resting state functional magnetic resonance imaging; sMRI, structural magnetic resonance imaging; ICA, independent component analysis; SVM, support vector machine; SNP, single nucleotide polymorphism; GM, gray matter; WM, white matter; FA, fractional anisotropy; PCA, principal component analysis; ν-SVR, ν support vector regression; EEG, electroencephalography; ALFF, amplitude of low-frequency fluctuation; mCCA, multi-set canonical correlation analysis; jICA, joint independent component analysis; SVM-RFE, support vector machine with recursive feature elimination; MEG, magnetoencephalography; Sg-ICA, spatial group independent component analysis; LDC, linear discriminant classifier; NBC, naïve Bayes classifier; HF, hippocampal formation; FLD, Fisher's linear discriminate; GSM, generalized sparse model; SRVS, sparse representation-based variable selection; ELM, extreme learning machine; LDA, linear discriminant analysis; LSVM, linear support vector machine; RSVM, radial basis function support vector machine; KNN, Amari k-nearest neighbor algorithm; GNB, Gaussian naïve Bayes; FAS, fractional anisotropy skeleton; SK, unweighted "simple" sum of kernels; MKL, multi-kernel learning; AV, prediction averaging; MV, majority voting; FEP, first episode psychosis; UHR, ultra-high risk; MDD, major depressive disorder; DTI, diffusion tensor imaging; TCP, transductive conformal predictor; LLD, late-life depression; MD, mean diffusivity; LR, logistic regression; TBSS, tract-based spatial statistics; ADHDI, attention-deficit/hyperactivity disorder, inattentive; ADHDC, attention-deficit/hyperactivity disorder, combined; LASSO, least absolute shrinkage and selection operator; NFS, no feature selection; RBF, radial basis function; NMF, nonnegative matrix factorization, CSF, cerebrospinal fluid; CT, cortical thickness; ReHo, regional homogeneity; FC, functional connectivity; MKL, multikernel learning; CFS, correlation-based feature selection; GPC, Gaussian process classifier; CORR, correlation; SIC, sparse inverse covariance; KAPPA, Patel's kappa.

promising to obtain higher accuracies than using single modality, it would be more powerful to include additional neuroimaging techniques. As far as we know, there is one study that combined rs-fMRI and MEG data. Çetin et al[112] differentiate SZ from HC with different classifiers and with an ensemble classifier. The best performance is provided by the combination of all by using the ensemble classifier (87.91%).

Besides sMRI, fMRI and genetic data, DTI and EEG can also be used to derive features for classification. Sui et al[34] combined rs-fMRI, sMRI and DTI (fractional anisotropy [FA]) with a fusion technique named as mCCA + jICA and used multiple type of classifiers. The most powerful prediction (>90% accuracy) can be accomplished using features from FA + gray matter (GM) via radial basis function support vector machine (RSVM). Sui et al[113] combined rs-fMRI, sMRI, and EEG and selected the features via mCCA + jICA, 2-sample *t* test and SVM with recursive feature elimination (SVM-RFE). They obtained 91% in training and 100% predication rate with combination of modalities.

To the best of our knowledge only 1 study combined 3 modalities (task fMRI, sMRI, and DTI) to classify patients with psychosis. Pettersson-Yeo et al[114] describe 4 integrative approaches to combine data obtained from task-related fMRI, sMRI, and DTI: (*a*) an unweighted sum of kernels, (*b*) multikernel learning, (*c*) prediction averaging, and (*d*) majority voting in order to classify ultra-high-risk (UHR) individuals for psychosis, first episode psychosis (FEP), and HC. The performance of the classifier (SVM) was 83.33% with combination of all modalities for differentiating FEP from UHR, and the best performance was 86.33% with combination of DTI and fMRI for differentiating FEP from UHR.

There are 3 studies that attempted to classify MDD patients. Schnyer et al[115] and Ota et al[116] combined sMRI and DTI data to differentiate MDD from SZ and HC, respectively. While Ota et al[116] used discriminant analysis for classification (72% for SZ, 88% for MDD), Schnyer et al[115] applied SVM to classify SZ and HC (70% for whole-brain FA, 74% for white matter FA). On the other hand, Schmaal et al[117] used combination of

task-related fMRI and sMRI of different types of patients with MDD and they classify them via binary Gaussian process classifier with the 62% accuracy for chronic and remitted MDD, 61% accuracy for chronic and gradually improved MDD, 44% accuracy for gradually improved and remitted MDD.

Classification of ADHD was studies by Qureshi et al,[118] Colby et al,[119] and Dai et al[120] using a combination of rs-fMRI and sMRI and by Anderson et al,[121] and Bohland et al[122] using rs-fMRI and sMRI, plus phenotypic data. Each study used a relatively large dataset and applied various types of classifiers and obtained different accuracies that change between 55% and 93% (please see Table 3). Moreover, Qureshi et al[118] reported that the classification accuracy for ADHD patients was 73.81%, 71.43%, and 76.16 based on sMRI, fMRI, and sMRI + fMRI data, respectively. However, this study had a smaller sample size in comparison to former studies. It should also be mentioned that these ADHD studies did not use a modality that has high temporal resolution such as EEG or MEG.

There are also some studies that used different tasks[75,123-125] or different features[126] of 1 modality as multiple modalities, but they are not included in Table 3 because different neuroimaging techniques were accepted as modalities in our overview.

Although there are some constraints such as small sample sizes and few N-way combination, the results of the studies given above are encouraging for using multimodal neuroimaging in classification of psychiatric diseases.

Conclusion and Future Directions

In recent years use of MN for the diagnosis of diseases gained momentum among researchers because of the limitations of unimodal neuroimaging techniques. In addition, the use of ML for early diagnoses, particularly psychiatric diseases, by neuroscientists is increasing. The present review aimed to provide an overview of published research based on MN and data fusion for classifying patients with psychiatric disorder/diseases and healthy controls, as well as an introduction to the types of data fusion approaches and ML techniques used for classification. Overall, the literature shows that MN improves the diagnostic prediction rate (accuracy) and provides more reliable classification of psychiatric diseases[107,111,114]; however, it is obvious that MN still has several challenges that need to be overcome, including population size, as with exception of a few meta-analyses, the other studies cited herein included small study populations. Furthermore, to date there remains the lack of an accurate ML technique for all applications; therefore, it may be useful to test various techniques with each dataset. In addition, more studies that use the N-way fusion model are needed, as the model could be helpful for obtaining more powerful results (eg, high accuracy) so the number of studies that combine multiple modalities should be increased. For instance, nonimaging predictors (such as age, gender, handedness, and cognitive ability) could be used as modality for the classification of diseases (eg, depression).

Author Contributions

EET contributed to literature search; writing; critically revised manuscript and gave final approval. BM contribued to literature search; writing; critically revised manuscript and gave final approval. NT contributed to conception and gave final approval. MKA is owner of the idea, contributed to literature search; writing; critically revised manuscript and gave final approval.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Emine Elif Tulay https://orcid.org/0000-0003-0150-5476

References

1. Liu S, Cai W, Liu S, et al. Multimodal neuroimaging computing: a review of the applications in neuropsychiatric disorders. *Brain Inform*. 2015;2:167-180.
2. Calhoun VD, Sui J. Multimodal fusion of brain imaging data: a key to finding the missing link(s) in complex mental illness. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2016;1:230-244.
3. Correa NM, Adali T, Li YO, Calhoun VD. Canonical correlation analysis for data fusion and group inferences. *IEEE Signal Process Mag*. 2010;27:39-50.
4. Wiecki TV, Poland J, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry. *Clin Psychol Sci*. 2015;3:378-399.
5. Nicolas-Alonso LF, Gomez-Gil J. Brain computer interfaces, a review. *Sensors*. 2012;12:1211-1279.
6. Malhi GS, Lagopoulos J. Making sense of neuroimaging in psychiatry. *Acta Psychiatr Scand*. 2008;117:100-117.
7. Calhoun VD, Adali T. Feature-based fusion of medical imaging data. *IEEE Trans Inf Technol Biomed*. 2009;13:711-720.
8. Bießmann F. *Data-Driven Analysis for Multimodal Neuroimaging* [master's thesis]. Berlin, Germany: Technischen Universität; 2012.
9. Mier W, Mier D. Advantages in functional imaging of the brain. *Front Hum Neurosci*. 2015;9:249.
10. Bunge SA, Kahn I. Cognition: an overview of neuroimaging techniques. In: Squire LR, ed. *Encyclopedia of Neuroscience*. New York, NY: Elsevier; 2009:1063-1067.
11. Bolognini N, Ro T. Transcranial magnetic stimulation: disrupting neural activity to alter and assess brain function. *J Neurosci*. 2010;30:9647-9650.
12. Rossi S, Hallett M, Rossini PM, Pascual-Leone A; Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol*. 2009;120:2008-2039.
13. Pascual-Marqui RD, Michel CM, Lehmann D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int J Psychophysiol*. 1994;18:49-65.

14. Pascual-Marqui RD, Esslen M, Kochi K, Lehmann D. Functional imaging with low resolution brain electromagnetic tomography (LORETA): a review. *Methods Find Exp Clin Pharmacol*. 2002;24(suppl C):91-95.
15. Phillips ML, Swartz HA. Critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. *Am J Psychiatry*. 2014;171:829-843.
16. Painold A, Faber PL, Milz P, et al. Brain electrical source imaging in manic and depressive episodes of bipolar disorder. *Bipolar Disord*. 2014;16:690-702.
17. Fonseka TM, MacQueen GM, Kennedy SH. Neuroimaging biomarkers as predictors of treatment outcome in major depressive disorder. *J Affect Disord*. 2018;233:21-35.
18. Yildirim A, Tureli D. Schizophrenia: a review of neuroimaging techniques and findings. *Eastern J Med*. 2015;20:1-6.
19. Bajouco M, Mota D, Coroa M, Caldeira S, Santos V, Madeira N. The quest for biomarkers in schizophrenia: from neuroimaging to machine learning. *Int J Clin Neurosci Ment Health*. 2017;4(suppl 3):S03.
20. Eugene AR, Masiak J. Electrophysiological neuroimaging using sLORETA comparing 100 schizophrenia patients to 48 patients with major depression. *Brain (Bacau)*. 2014;5:16-25.
21. Hughes JR, John ER. Conventional and quantitative electroencephalography in psychiatry. *J Neuropsychiatry Clin Neurosci*. 1999;11:190-208.
22. Rosa MJ, Daunizeau J, Friston KJ. EEG-fMRI integration: a critical review of biophysical modeling and data analysis approaches. *J Integr Neurosci*. 2010;9:453-476.
23. Laufs H. A personalized history of EEG-fMRI integration. *Neuroimage*. 2012;62:1056-1067.
24. Bridwell D, Calhoun V. Fusing concurrent EEG and fMRI intrinsic networks. In: Supek S, Aine CJ, eds. *Magnetoencephalography: From Signals to Dynamic Cortical Networks*. Berlin, Germany: Springer-Verlag; 2014:213-235.
25. Rosenkranz K, Lemieux L. Present and future of simultaneous EEG-fMRI. *MAGMA*. 2010;23:309-316.
26. Uludağ K, Roebroeck A. General overview on the merits of multimodal neuroimaging data fusion. *Neuroimage*. 2014;102(pt 1):3-10.
27. Lahat D, Adali T, Jutten C. Multimodal data fusion: an overview of methods, challenges, and prospects. *Proc IEEE*. 2015;103:1449-1477.
28. Biessmann F, Plis S, Meinecke FC, Eichele T, Müller KR. Analysis of multimodal neuroimaging data. *IEEE Rev Biomed Eng*. 2011;4:26-58.
29. Sui J, Pearlson G, Caprihan A, et al. Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA+ joint ICA model. *Neuroimage*. 2011;57:839-855.
30. Sui J, Yu Q, He H, Pearlson GD, Calhoun VD. A selective review of multimodal fusion methods in schizophrenia. *Front Hum Neurosci*. 2012a;6:27.
31. Sui J, Huster R, Yu Q, Segall JM, Calhoun VD. Function-structure associations of the brain: evidence from multimodal connectivity and covariance studies. *Neuroimage*. 2014a;102(pt 1):11-23.
32. Schultz CC, Fusar-Poli P, Wagner G, et al. Multimodal functional and structural imaging investigations in psychosis research. *Eur Arch Psychiatry Clin Neurosci*. 2012;262 (suppl 2):S97-S106.
33. Ahn S, Jun SC. Multi-modal integration of EEG-fNIRS for brain-computer interfaces—current limitations and future directions. *Front Hum Neurosci*. 2017;11:503.
34. Sui J, He H, Pearlson GD, et al. Three-way (N-way) fusion of brain imaging data based on mCCA+jICA and its application to discriminating schizophrenia. *Neuroimage*. 2013;66:119-132.
35. 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.
36. Lahat D, Adali T, Jutten C. Challenges in multimodal data fusion. Paper presented at: 22nd European Signal Processing Conference (EUSIPCO-2014); September 1-5, 2014; Lisbon, Portugal.
37. Huster RJ, Debener S, Eichele T, Hermann CS. Methods for simultaneous EEG-fMRI: an introductory review. *J Neurosci*. 2012;32:6053-6060.
38. He B, Liu Z. Multimodal functional neuroimaging: integrating functional MRI and EEG/MEG. *IEEE Rev Biomed Eng*. 2008;1:23-40. doi:10.1109/RBME.2008.2008233.
39. Valdes-Sosa PA, Sanchez-Bornot JM, Sotero RC, et al. Model driven EEG/fMRI fusion of brain oscillations. *Hum Brain Mapp*. 2009;30:2701-2721.
40. Gilron R, Rosenblatt J, Koyejo O, Poldrack RA, Mukamel R. What's in a pattern? Examining the type of signal multivariate analysis uncovers at the group level. *Neuroimage*. 2017;146:113-120.
41. Michael AM, King MD, Ehrlich S, et al. A data-driven investigation of gray matter-function correlations in schizophrenia during a working memory task. *Front Hum Neurosci*. 2011;5:71.
42. Haller S, Xekardaki A, Delaloye C, et al. Combined analysis of grey matter voxel-based morphometry and white matter tract-based spatial statistics in late-life bipolar disorder. *J Psychiatry Neurosci*. 2011;36:391-401.
43. Chen Z, Cui L, Li M, et al. Voxel based morphometric and diffusion tensor imaging analysis in male bipolar patients with first-episode mania. *Prog Neuropsychopharmacol Biol Psychiatry*. 2012;36:231-238.
44. Kawaguchi H, Shimada H, Kodaka F, et al. Principal component analysis of multimodal neuromelanin MRI and dopamine transporter PET data provides a specific metric for the nigral dopaminergic neuronal density. *PLoS One*. 2016;11:e0151191.
45. Stephen JM, Coffman BA, Jung RE, Bustillo JR, Aine CJ, Calhoun VD. Using joint ICA to link function and structure using MEG and DTI in schizophrenia. *Neuroimage*. 2013;83:418-430.
46. Jagannathan K, Calhoun VD, Gelernter J, et al. Genetic associations of brain structural networks in schizophrenia: a preliminary study. *Biol Psychiatry*. 2010;68:657-666.
47. Meda S, Jagannathan K, Gelernter J, et al. A pilot multivariate parallel ICA study to investigate differential linkage between neural networks and genetic profiles in schizophrenia. *Neuroimage*. 2010;53:1007-1015.
48. Bießmann F, Meinecke F, Gretton A, et al. Temporal kernel CCA and its application in multimodal neuronal data analysis. *Machine Learning*. 2010;79:5-27.
49. Krishnan A, Williams L, McIntosh A, Abdi H. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage*. 2011;56:455-475.

50. Groves AR, Beckmann CF, Smith SM, Woolrich MW. Linked independent component analysis for multimodal data fusion. *Neuroimage*. 2011;54:2198-2217.
51. Mwangi B, Tian T, Soares JC. A review of feature reduction techniques in neuroimaging. *Neuroinformatics*. 2014b;12:229-244.
52. Kim DI, Sui J, Rachakonda S, et al. Identification of imaging biomarkers in schizophrenia: a coefficient-constrained independent component analysis of the mind multi-site schizophrenia study. *Neuroinformatics*. 2010;8:213-229.
53. Sui J, He H, Yu Q, et al. Combination of resting state fMRI, DTI, and sMRI data to discriminate schizophrenia by N-way MCCA + jICA. *Front Hum Neurosci*. 2013b;7:235.
54. Abrol A, Rashid B, Rachakonda S, Damaraju E, Calhoun VD. Schizophrenia shows disrupted links between brain volume and dynamic functional connectivity. *Front Neurosci*. 2017;11:624.
55. Kim SG, Jung WH, Kim SN, Jang JH, Kwon JS. Alterations of gray and white matter networks in patients with obsessive-compulsive disorder: a multimodal fusion analysis of structural MRI and DTI using mCCA + jICA. *PLoS One*. 2015;10:e0127118.
56. Sui J, Adali T, Yu Q, Chen J, Calhoun VD. A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci Methods*. 2012b;204:68-81.
57. Mitchell TM, Hutchinson R, Niculescu RS, et al. Learning to decode cognitive states from brain images. *Machine Learn*. 2004;57:145-175.
58. Friston KJ. Modalities, modes, and models in functional neuroimaging. *Science*. 2009;326:399-403.
59. Bzdok D. Classical statistics and statistical learning in imaging neuroscience. *Front Neurosci*. 2017;11:543.
60. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. *Neuroimage*. 2017;145(pt B):137-165.
61. Glaser JI, Chowdhury RH, Perich MG, Miller LE, Kording KP. Machine learning for neural decoding. arXiv:1708.00909 [q-bio.NC]. https://arxiv.org/ftp/arxiv/papers/1708/1708.00909.pdf 2017. Accessed May 28, 2018.
62. Lemm S, Blankertz B, Dickhaus T, Muller KR. Introduction to machine learning for brain imaging. *Neuroimage*. 2011;56:387-399.
63. Zarogianni E, Moorhead TW, Lawrie SM. Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. *Neuroimage Clin*. 2013;3:279-289.
64. Bray S, Chang C, Hoeft F. Applications of multivariate pattern classification analyses in developmental neuroimaging of healthy and clinical populations. *Front Hum Neurosci*. 2009;3:32.
65. Orrù G, Pettersson-Yeo W, Marquand A, 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.
66. Vieira S, Pinaya WH, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. *Neurosci Biobehav Rev*. 2017;74(pt A):58-75.
67. Ryali S, Supekar K, Abrams D, Menon V. Sparse logistic regression for whole-brain classification of fMRI data. *Neuroimage*. 2010;51:752-764.
68. Zhong M, Lotte F, Girolami M, et al. Classifying EEG for brain computer interfaces using Gaussian processes. *Pattern Recognit Lett*. 2008;29:354-359.
69. Prichep LS, John ER. QEEG profiles of psychiatric disorders. *Brain Topogr*. 1992;4:249-257.
70. Pardo PJ, Georgopoulos AP, Kenny JT, Stuve TA, Findling RL, Schulz SC. Classification of adolescent psychotic disorders using linear discriminant analysis. *Schizophr Res*. 2006;87:297-306.
71. Serra A, Galdi P, Tagliaferri R. Machine learning for bioinformatics and neuroimaging [published online February 22, 2018]. *WIREs Data Mining Knowl Discov*. doi:10.1002/widm.1248.
72. Salvador R, Radua J, Canales-Rodríguez EJ, et al. Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One*. 2017;12:e0175683.
73. Bhaumik R, Jenkins L, Gowins J, et al. Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity. *Neuroimage Clin*. 2016;16:390-398.
74. Lu X, Yang Y, Wu F, et al. Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine (Baltimore)*. 2016;95:e3973.
75. Cetin M, Stephen J, Calhoun VD. Sensory load hierarchy-based classification of schizophrenia patients. Paper presented at: IEEE International Conference on Image Processing (ICIP); September 27-30, 2015; Quebec City, Quebec, Canada.
76. Haller S, Lovblad KO, Giannakopoulos P, Van De Ville D. Multivariate pattern recognition for diagnosis and prognosis in clinical neuroimaging: state of the art, current challenges and future trends. *Brain Topogr*. 2014;27:329-337.
77. Klöppel S, Abdulkadir A, Jack CR Jr, Koutsouleris N, Mourão-Miranda J, Vemuri P. Diagnostic neuroimaging across diseases. *Neuroimage*. 2012;61:457-463.
78. Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*. 2009;45(1 suppl):S199-S209.
79. Chu C, Hsu AL, Chou KH, Bandettini P, Lin C; Alzheimer's Disease Neuroimaging Initiative. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *Neuroimage*. 2012;60:59-70.
80. Schnack HG, Kahn RS. Detecting neuroimaging biomarkers for psychiatric disorders: sample size matters. *Front Psychiatry*. 2016;7:50.
81. O'Halloran R, Kopell BH, Sprooten E, Goodman WK, Frangou S. Multimodal neuroimaging-informed clinical applications in neuropsychiatric disorders. *Front Psychiatry*. 2016;7:63.
82. Koutsouleris N, Davatzikos C, Borgwardt S, et al. Accelerated brain aging in schizophrenia and beyond: a neuroanatomical marker of psychiatric disorders. *Schizophr Bull*. 2014;40:1140-1153.
83. Castellanos FX, Di Martino A, Craddock RC, Mehta AD, Milham MP. Clinical applications of the functional connectome. *Neuroimage*. 2013;80:527-540.
84. Patel MJ, Khalaf A, Aizenstein HJ. Studying depression using imaging and machine learning methods. *Neuroimage Clin*. 2016;10:115-123.
85. Kambeitz J, Cabral C, Sacchet MD, et al. Detecting neuroimaging biomarkers for depression: a meta-analysis of multivariate pattern recognition studies. *Biol Psychiatry*. 2017;82:330-338.
86. Erguzel T, Ozekes S, Bayram A, et al. Classification of major depressive disorder subjects using Pre-rTMS electroencephalography data with support vector machine approach. Paper

presented at: IEEE, Science and Information Conference (SAI); August 27-29, 2014; London, UK.
87. Erguzel TT, Ozekes S, Gultekin S, Tarhan N, Hizli Sayar G, Bayram A. Neural network based response prediction of rTMS in major depressive disorder using QEEG cordance. *Psychiatr Investig*. 2015;12:61-65.
88. Erguzel TT, Sayar GH, Tarhan N. Artificial intelligence approach to classify unipolar and bipolar depressive disorders. *Neural Comput Appl*. 2016;27:1607-1616.
89. Kotsilieris T, Pintelas E, Livieris IE, Pintelas P. Reviewing machine learning techniques for predicting anxiety disorders 2018. Report no. TR01-18. http://nemertes.lis.upatras.gr/jspui/bitstream/10889/10981/6/TR01-18.pdf. Accessed May 28, 2018.
90. Frick A, Gingnell M, Marquand AF, et al. Classifying social anxiety disorder using multivoxel pattern analyses of brain function and structure. *Behav Brain Res*. 2014;259:330-335.
91. Isobe M, Miyata J, Hazama M, Fukuyama H3, Murai T, Takahashi H. Multimodal neuroimaging as a window into the pathological physiology of schizophrenia: current trends and issues. *Neurosci Res*. 2016;102:29-38.
92. Kambeitz J, Kambeitz-Ilankovic L, Leucht S, et al. Detecting neuroimaging biomarkers for schizophrenia: a meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology*. 2015;40:1742-1751.
93. Demirci O, Clark VP, Magnotta VA, et al. A review of challenges in the use of fMRI for disease classification/characterization and a projection pursuit application from a multi-site fMRI schizophrenia study. *Brain Imaging Behav*. 2008;2: 147-226.
94. Zhu C, Zang YF, Cao QJ, et al. Fisher discriminative analysis of resting-state brain function for attention-deficit/hyperactivity disorder. *Neuroimage*. 2008;40:110-120.
95. Chabot RJ, Serfontein G. Quantitative electroencephalographic profiles of children with attention deficit disorder. *Biol Psychiatry*. 1996;40:951-963.
96. Chabot RJ, Merkin H, Wood LM, Davenport TL, Serfontein G. Sensitivity and specificity of QEEG in children with attention deficit or specific developmental learning disorders. *Clin Electroencephalogr*. 1996;27:26-34.
97. Chabot RJ, Orgill AA, Crawford G, Harris MJ, Serfontein G. Behavioral and electrophysiologic predictors of treatment response to stimulants in children with attention disorders. *J Child Neurol*. 1999;14:343-351.
98. Uddin LQ, Dajani DR, Voorhies W, Bednarz H, Kana RK. Progress and roadblocks in the search for brain-based biomarkers of autism and attention-deficit/hyperactivity disorder. *Transl Psychiatry*. 2017;7:e1218. doi:10.1038/tp.2017.164.
99. Dosenbach NUF, Nardos B, Cohen AL, et al. Prediction of individual brain maturity using fMRI. *Science*. 2010;329: 1358-1361.
100. Lynch CJ, Uddin LQ, Supekar K, Khouzam A, Phillips J, Menon V. Default mode network in childhood autism: posteromedial cortex heterogeneity and relationship with social deficits. *Biol Psychiatry*. 2013;74:212-219.
101. Mwangi B, Matthews K, Steele JD. Prediction of illness severity in patients with major depression using structural MR brain scans. *J Magn Reson Imaging*. 2012;35:64-71.
102. Meng X, Jiang R, Lin D, et al. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage*. 2017;145(pt B):218-229.
103. Mwangi B, Soares JC, Hasan KM. Visualization and unsupervised predictive clustering of high-dimensional multimodal neuroimaging data. *J Neurosci Methods*. 2014a;236:19-25.
104. Zeng LL, Shen H, Liu L, Hu D. Unsupervised classification of major depression using functional connectivity MRI. *Hum Brain Mapp*. 2014;35:1630-1641.
105. John ER, Prichep LS, Almas M. Subtyping of psychiatric patients by cluster analysis of QEEG. *Brain Topogr*. 1992;4:321-326.
106. Qureshi MNI, Oh J, Cho D, Jo HJ, Lee B. Multimodal discrimination of schizophrenia using hybrid weighted feature concatenation of brain functional connectivity and anatomical features with an extreme learning machine. *Front Neuroinform*. 2017a;11:59.
107. Cabral C, Kambeitz-Ilankovic L, Kambeitz J, et al. Classifying schizophrenia using multimodal multivariate pattern recognition analysis: evaluating the impact of individual clinical profiles on the neurodiagnostic performance. *Schizophr Bull*. 2016;42(suppl 1):S110-S117.
108. Yang H, He H, Zhong J. Multimodal MRI characterization of schizophrenia: a discriminative analysis. *Lancet*. 2016;388:S36.
109. Ford J, Shen L, Makedon F, Flashman LA, Saykin AJ. A combined structural-functional classification of schizophrenia using hippocampal volume plus fMRI activation. Paper presented at: Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society/Engineering in Medicine and Biology; October 23-26, 2002; Houston, TX.
110. Cao L, Guo S, Xue Z, et al. Aberrant functional connectivity for diagnosis of major depressive disorder: a discriminant analysis. *Psychiatry Clin Neurosci*. 2014;68:110-119.
111. Yang H, Liu J, Sui J, Pearlson G, Calhoun VD. A hybrid machine learning method for fusing fMRI and genetic data: combining both improves classification of schizophrenia. *Front Hum Neurosci*. 2010;4:192.
112. Cetin MS, Houck JM, Rashid B, et al. Multimodal classification of schizophrenia patients with MEG and fMRI data using static and dynamic connectivity measures. *Front Neurosci*. 2016;10:466.
113. Sui J, Castro E, Hao H, et al. Combination of FMRI-SMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection. Paper presented at: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 29-30, 2014; Chicago, IL.
114. Pettersson-Yeo W, Benetti S, Marquand AF, et al. An empirical comparison of different approaches for combining multimodal neuroimaging data with support vector machine. *Front Neurosci*. 2014;8:189.
115. Schnyer DM, Clasen PC, Gonzalez C, Beevers CG. Evaluating the diagnostic utility of applying a machine learning algorithm to diffusion tensor MRI measures in individuals with major depressive disorder. *Psychiatr Res Neuroimaging*. 2017;264: 1-9.
116. Ota M, Ishikawa M, Sato N, et al. Discrimination between schizophrenia and major depressive disorder by magnetic resonance imaging of the female brain. *J Psychiatr Res*. 2013;47:1383-1388.
117. 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.

118. Qureshi MNI, Oh J, Min B, Jo HJ, Lee B. Multi-modal, multi-measure, and multi-class discrimination of ADHD with hierarchical feature extraction and extreme learning machine using structural and functional brain MRI. *Front Hum Neurosci*. 2017b;11:157.
119. Colby JB, Rudie JD, Brown JA, Douglas PK, Cohen MS, Shehzad Z. Insights into multimodal imaging classification of ADHD. *Front Syst Neurosci*. 2012;6:59.
120. Dai D, Wang J, Hua J, He H. Classification of ADHD children through multimodal magnetic resonance imaging. *Front Syst Neurosci*. 2012;6:63.
121. Anderson A, Douglas PK, Kerr WT, et al. Non-negative matrix factorization of multimodal MRI, fMRI and phenotypic data reveals differential changes in default mode subnetworks in ADHD. *Neuroimage*. 2014;102(pt 1):207-219.
122. Bohland JW, Saperstein S, Pereira F, Rapin J, Grady L. Network, anatomical, and non-imaging measures for the prediction of ADHD diagnosis in individual subjects. *Front Syst Neurosci*. 2012;6:78.
123. Levin-Schwartz Y, Calhoun VD, Adali T. Quantifying the interaction and contribution of multiple datasets in fusion: application to the detection of schizophrenia. *IEEE Tran Med Imaging*. 2017;36:1385-1395.
124. Ramasubbu R, Brown MRG, Cortese F, et al. Accuracy of automated classification of major depressive disorder as a function of symptom severity. *Neuroimage Clin*. 2016;12:320-331.
125. Michael AM, Calhoun VD, Andreasen NC, Baum SA. A method to classify schizophrenia using inter-task spatial correlations of functional brain images. *Conf Proc IEEE Eng Med Biol Soc*. 2008;2008:5510-5513.
126. Cheng W, Ji X, Zhang J, Feng J. Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. *Front Syst Neurosci*. 2012;6:58.