



Advances in Using MRI to Estimate the Risk of Future Outcomes in Mental Health - Are We Getting There?

Aleix Solanes^{1,2*} and Joaquim Radua^{1,3,4}

¹ Imaging of Mood- and Anxiety-Related Disorders (IMARD) Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain, ² Department of Psychiatry and Forensic Medicine, School of Medicine, Autonomous University of Barcelona, Barcelona, Spain, ³ Early Psychosis: Interventions and Clinical-detection Lab, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, ⁴ Department of Clinical Neuroscience, Stockholm Health Care Services, Stockholm County Council, Karolinska Institutet, Stockholm, Sweden

Keywords: neuroimaging, machine-learning (ML), risk estimating, schizophrenia, magnetic resonance imaging (MRI)

INTRODUCTION

Schizophrenia is a mental disorder among the leading disabling conditions worldwide (1). It affects approximately 20 million people and increases 2–3 times the probability of dying early and decreases the life expectancy by about 20 years (2). Its onset is typically in adolescence or early adulthood (3). Relevantly, prolonged untreated psychosis leads to poorer outcomes (4), so detecting the disorder early could considerably improve their lives.

Neuroimaging studies have found that individuals with schizophrenia have smaller volumes in the hippocampus, amygdala, thalamus, nucleus accumbens, intracranial space, and larger pallidum and ventricle volumes (5–9).

Traditionally, most neuroimaging studies in psychiatry relied on mass-univariate statistical approaches. For example, techniques like voxel-based morphometry (VBM) let researchers assess voxel-wise differences in regional volume or tissue composition based on estimates of tissue probability. These approaches are helpful to detect group differences. Still, they cannot yield predictions (e.g., diagnosis, outcome, etc.) at the individual level. In recent years, the interest in machine-learning methods in neuroimaging has increased due to its capability to handle high-dimensional data and perform predictions at a single-subject level.

The use of machine learning algorithms such as Support-Vector Machines (10), regularized regression (11, 12), Random Forest (13), or more recently, Deep Learning (14) to different neuroimaging modalities has expanded the possibilities of brain imaging data much beyond the traditional case-control group comparisons.

For example, many researchers aimed to use MRI-based data to detect mental disorders. Most focused on classifying each MRI scan as being from a patient or a healthy control (15–17). Still, others attempted to classify each MRI scan as being from a patient with one or another mental disorder (18–20). A recent systematic review summarized the classification performance between patients with schizophrenia and healthy control reported by various studies. High-performance prediction accuracy of >70% was reported in 40 of 41 studies using structural MRI, 35 of 40 using functional MRI scans, and 5 of 5 using diffusion-weight MRI (21). While these prediction rates may seem impressive, most of these diagnostic tools have not been integrated into clinical practice.

Less machine learning research has been done in detecting subjects at high risk of developing future outcomes in schizophrenia (22, 23). This is unfortunate because early detection could delay or even prevent severe future consequences (24). Some studies used clinical data, such as the presence of manic and negative symptoms (25–27), the diagnosis at onset together with other

OPEN ACCESS

Edited by:

Stefan Borgwardt,
University of Lübeck, Germany

Reviewed by:

Alexandra Korda,
University Medical Center
Schleswig-Holstein, Germany
Gemma C. Monté-Rubio,
University of Barcelona, Spain

*Correspondence:

Aleix Solanes
solanes@clinic.cat

Specialty section:

This article was submitted to
Neuroimaging and Stimulation,
a section of the journal
Frontiers in Psychiatry

Received: 30 November 2021

Accepted: 16 March 2022

Published: 12 April 2022

Citation:

Solanes A and Radua J (2022)
Advances in Using MRI to Estimate
the Risk of Future Outcomes in Mental
Health - Are We Getting There?
Front. Psychiatry 13:fpsyt-13-826111.
doi: 10.3389/fpsy.2022.826111

sociodemographic and clinical scales information (28, 29), or drugs use (28, 30, 31). Others used biological information such as blood-based biomarkers (32, 33), genetics data (34–36), or the combination of both clinical and biological data (37).

Finally, the use of MRI data to estimate the risk of different outcomes in schizophrenia has been even less explored. Only a few studies have taken this path, using as predictor variables the changes in brain volume during the 1st year after a first episode of psychosis (FEP) (38), brain gray matter MRI (39, 40), or surface-based data (40).

While results are still humble, we believe that these efforts are building a base on top of which future research will create valuable tools that will help the clinician. For example, these tools could help detect subjects at risk or predict the response to different treatments with the final aim to improve the patient wellbeing (41, 42).

The following section will describe which machine learning algorithms are currently used in neuroimaging. Then we will review which are the most common pitfalls and errors that we believe researchers should avoid. Finally, we will set our sights on the future, checking some of the promises of the latest algorithms and neuroimaging techniques.

COMMON MACHINE LEARNING ALGORITHMS

The field of machine learning englobes an extensive list of algorithms, each with its strengths and weaknesses. Each of these algorithms should be tuned to obtain a trade-off between the model not fitting the data well enough, which is called underfitting, and excessive fitting the training data, called overfitting (see **Figure 1**). An underfitted model fails to capture the relationship between the MRI data and the outcome, for what it performs poorly even on the training data. Overfitting occurs when the model finds relationships between the MRI data and the outcome that are only based on particular, random details of the training data, and thus its performance is poor on new data. Over the vast selection of algorithms, we will review only the more common.

Logistic Regression

Logistic Regression (LR) is a statistical model that fits the data to a logit or logistic function, an S-shaped sigmoid function. It calculates a binary dependent variable from one or more independent variables. Since neuroimaging results in a huge number of variables and many of them may be highly correlated, a technique called regularization is often used to reduce the complexity. Among the most used regularization algorithms in linear regression, we can find Least Absolute Shrinkage and Selection Operator (LASSO) (12), which eliminates non-important variables from the final model that do not contribute much to the prediction task, or Ridge (11), which instead of removing the variables, applies a penalty to some values that result in near-zero coefficients. A combination of both Ridge and LASSO also exists, which is called Elastic Net (43). This algorithm

eliminates the coefficients of some variables and also reduces others close to zero.

Despite being a simple algorithm, it has been used to predict subjects at ultra-high risk and even performed better than more complex algorithms (44). Elastic Net was used to detect both functional and structural brain alterations in female patients with schizophrenia (45). Salvador et al. evaluated the discriminative power of some of the most commonly used algorithms on sMRI for prediction in psychosis. They systematically compared the different algorithms, and among them, Ridge, LASSO, and Elastic Net classifiers performed similarly, if not slightly better, than other classifiers (46).

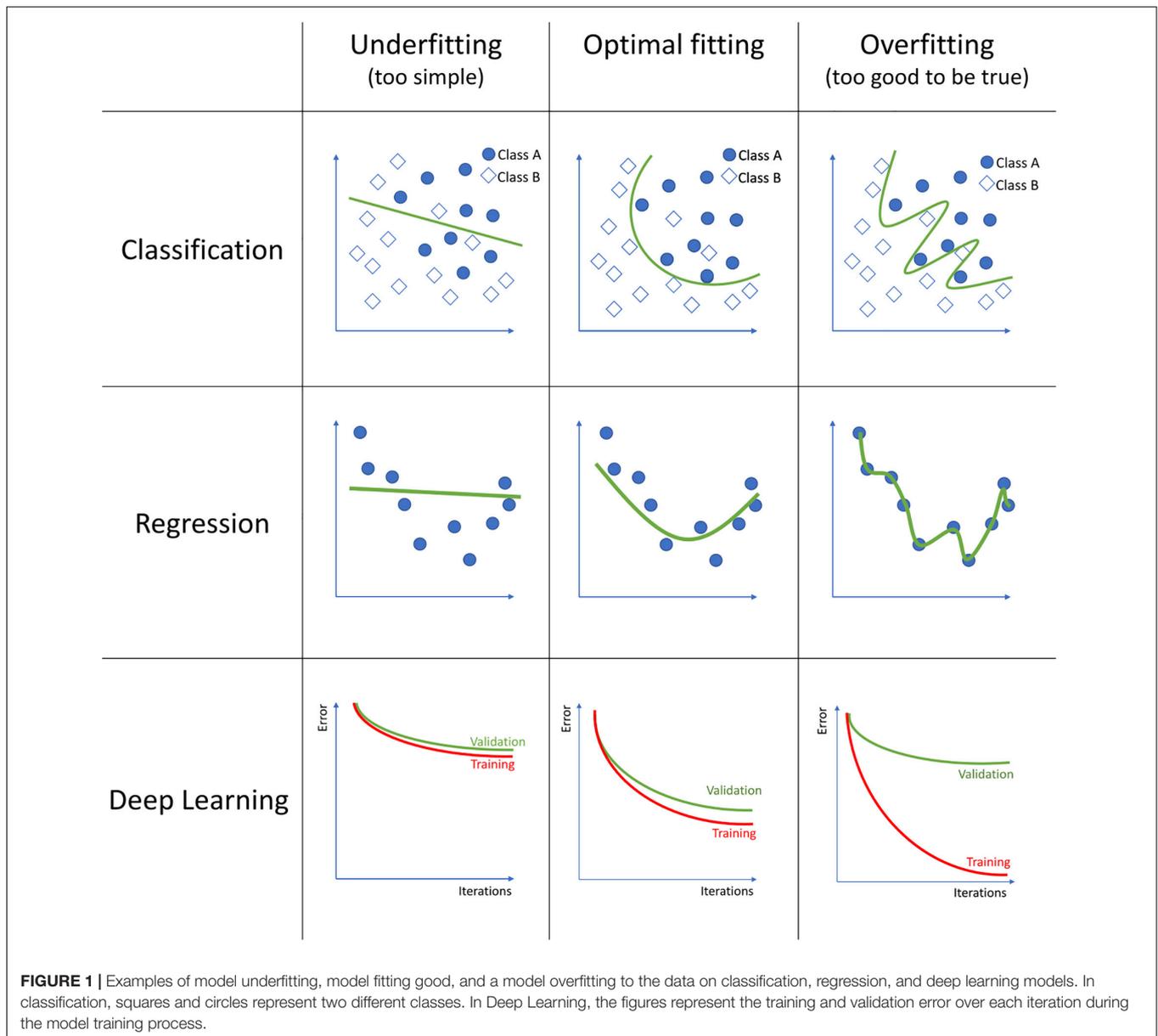
Support Vector Machines

Support Vector Machine (SVM) is a supervised discriminative classification method that creates hyperplanes for optimally separating the data into different groups that can be used for both classification or regression tasks. One strength of the algorithm is the ability to perform non-linear classifications. This is done by taking low dimensional input space data and convert them into a higher dimensional input space so that a hyperplane can separate the different classes. To improve the generalization of the model and avoid overfitting, a trade-off between maximizing the margin between the classes and minimizing the number of misclassifications has to be considered.

SVM algorithms emerged as powerful tools for finding objective neuroanatomical biomarkers since they can quite effectively handle high-dimensional data, consider inter-regional correlations between different brain regions, and make inferences at a single-subject level with a decent classification result (47).

It is probably one of the most commonly used machine learning classification algorithms in neuroimaging data. A common approach has been the use of SVM to classify subjects at ultra-high risk for psychosis from healthy controls (48, 49, 50). Due to its capability to deal with high-dimensional data, it has also been applied to classify between recent-onset depression and recent-onset psychosis using both neuroanatomical information and clinical data (51), to find neurocognitive subtypes based on cognitive performance and neurocognitive alterations in recent onset psychosis (52, 53), to identify schizophrenia patients based on subcortical regions (54) or functional network connectivity data (55). A multimodal approach combining structural MRI, diffusion tensor imaging, and resting-state functional MRI data was tested to classify patients with chronic schizophrenia vs. patients with FEP comparing different algorithms such as Random Forest (RF), LR, Linear Discriminant Analysis (LDA), and K-Nearest Neighbor classification (KNN), and SVM, resulting in the latter as the best performing one (56). Steardo et al. in a recent systematic review, analyzed 22 studies using SVM on fMRI as biomarkers to classify between schizophrenia patients and controls, where 19 studies reported a promising >70% accuracy (57).

The popularity of SVM and the trade-off between performance and simplicity are its main strengths. It is essential to consider that its performance is highly dependent on the hyperparameters chosen, and these are usually tuned using a grid search method (58).



Linear Discriminant Analysis

Similar to SVM, linear discriminant analysis involves transforming the dimensionality of the data. However, in LDA, the data is projected into a lower-dimensional space where the different data groups can be maximally separated using a non-linear kernel (59).

It has been used on classification tasks such as Recent-Onset Schizophrenia (ROS) vs. HC (60), patients with FEP vs. HC (61–63), or patients with SZ vs. HC (64–67) with accuracies over 70%. Winterburn et al. used three independent datasets to validate the discriminative power of LDA to classify patients with SZ from HC, using different neuroimaging data. They compared cortical thickness data, ravens maps, and modulated VBM, and resulted that using their larger dataset, the accuracy was slightly lower compared to previous lower sample dataset articles (68).

The main strength of LDA is the reduction of the overfitting problem and computational cost by reducing the dimensional space. Despite that, its major drawback is that it requires assuming that the covariance matrix in the groups of data is identical, which is rare in real-world data.

Decision Trees and Random Forest

Decision trees are non-parametric supervised learning methods used for both classification and regression. They can predict values by learning simple decision rules inferred from the data features. These algorithms tend to overfit, and so, to not generalize well to new data. To overcome this limitation, a variation on this algorithm exists called Random Forests (RF). It is simply a collection of Decision Trees whose results are aggregated into one single final result in the end. RF incorporates

interactions between predictors in the model, detecting both linear and non-linear relationships.

It has been used to classify groups, such as childhood-onset schizophrenia patients and healthy controls (69), or schizophrenia, bipolar disorder, and healthy controls (70, 71).

It is an algorithm that generally provides high accuracy and a balance between bias-variance trade-off. Its major drawbacks are that it tends to be computationally intense on large datasets. It can be difficult to interpret the results, as it is difficult to analyze all coefficients.

Artificial Neural Networks and Deep Learning

Artificial Neural Networks (ANN) are a family of machine learning algorithms inspired by the brain's biological functioning. Like in our brain, these algorithms have neurons that receive a signal, process it, send a signal to the following connected neuron, and so on until a final result is obtained. To adjust the learning capabilities of the model, each neuron and synapse can have weights to increase or decrease the strength of the signal. Neurons are aggregated into layers, and when the number of layers increases, the algorithm is then known as Deep Learning. These advanced models can extract complex latent features from minimally preprocessed original data through non-linear transformations. To avoid overfitting, a method called Dropout exists. Dropout is a regularization method that randomly ignores or "drops out" some layer nodes, which is similar to adding noise to the training process. This improves the generalization of the model and reduces overfitting.

Deep Learning is often used to classify, predict values or even detect or segment regions in the brain.

Deep Learning is a vast topic that has increased the performance in some classification/prediction problems due to finding complex patterns on highly complex data. Despite being widely used to perform automatic tumor (72) or multiple sclerosis lesion detection (73) in brain MR images, it is still not extensively used in mental health disorder detection or risk-estimation of outcomes.

To overcome the black-box problem in artificial neural networks, different approaches are being developed, such as creating heat maps using Layer-Wise Propagation to identify the more important features involved in the algorithm decision (74, 75).

Libraries and Tools

Machine learning analyses involve choosing both the algorithm and its parameters. The research on machine learning applied to neuroimaging wills to boost the algorithm's performance. Still, some research groups decided to make their models available to the broad research community as open-source tools to let others use these improvements. This also enables researchers without coding backgrounds to perform machine learning analyses and expand the applicability and the validation of these tools. Some of these open-source tools focused on MRI machine learning are listed here:

NeuroMiner¹: It is a free Matlab toolbox developed by Nikolaos Koutsoleris with support from the PRONIA project that provides machine learning methods for analyzing heterogeneous data, such as clinical, structural, genetic, and functional neuroimaging data. It is designed to be easy to use, and no coding skills are required. Different MRI preprocessing steps are integrated into the software pipeline. It can create models to be applied to new data. It is designed to be flexible across different machine learning methods and data types, and it lets model sharing through a collaborative model library.

MRIpredict²: It is a free GUI-based tool and also an R library maintained by the IMARD Group at IDIBAPS, Hospital Clínic de Barcelona. Like NeuroMiner, it provides machine learning methods to predict diagnosis from structural MRI data and clinical information. It can create models to be applied to new data and makes it easy to analyze which variables are useful for prediction or classification. It can handle multisite data by ComBat harmonization (76), perform techniques like ensemble learning to improve the robustness of the results, or multiple imputations to handle missing fields data. In addition, it can also conduct other types of predictions and estimations, such as the risk of developing a specific outcome based on the time to the event (survival analyses). It also can be used as an R library.

Pronto³: It is also a free GUI-based Matlab developed by a team of researchers led by Prof. Janaina Mourao-Miranda. It provides various tools and algorithms for easily conducting machine learning analysis on neuroimaging data. The software lets the user specify the steps of the analysis as a batch job, choosing among the different functions of the software per each step.

Despite all the different algorithms, systematic comparisons between standard algorithms in neuroimaging pointed out that the differences in predictive performance are more attributable to differences in feature vectors than to the algorithm by itself (47, 77).

COMMON PITFALLS AND ERRORS

Reproducibility and Replicability

While the following definitions are not universal, reproducibility commonly refers to obtaining (approximately) the same results of an article using the same data and experimental procedures used in that investigation. In contrast, replicability commonly refers to obtaining consistent results using new data.

Neuroimaging studies typically involve many possible choices during imaging quality control, preprocessing, or statistical analyses. In addition, in machine learning techniques, each algorithm may have several parameters that authors can set, e.g., to find the best method for their analyses.

To achieve reproducibility, all brain imaging and machine learning choices should be detailed and shared by the authors. Ideally, papers should report all details and choices made during the study. In addition, making the models, code, and data

¹<http://proniapredictors.eu/neurominer/>

²<https://mripredict.com>

³<http://www.mlnl.cs.ucl.ac.uk/pronto/>

publicly available should be standard practice to let third parties examine the analysis in-depth and let others reproduce the whole process applied. That said, data may often be not sharable for privacy reasons.

In contrast, to achieve replicability, independent replication studies should be encouraged. Unfortunately, one common pitfall that prevents the applicability of many of these complex algorithms is their low replicability, their lack of generalizability. Many are developed on small sample sizes or single-site datasets because acquiring neuroimaging data is time-consuming and costly. However, while models made with small datasets or single-site data may seem to perform well, this performance is based on overfitting and, thus, the models fail when applied to new data. Like in human learning, the more examples an algorithm sees, the more it will learn to extrapolate those results to new samples (78). For example, recently, some authors evaluated some published models using clinical and neuropsychological data to predict the transition to psychosis in subjects at clinical high-risk for psychosis. When applying those published methods to a new sample, previously published models failed to predict or showed a poor accuracy (79, 80).

It is not uncommon in machine learning to test a battery of classifiers and report the one with the best accuracy. However, science should avoid studies based on cherry-picking, meaning taking only results advantageous for our research question. In some cases, this can be helpful, but this may easily lead to data torturing.

Bias

Creating and validating machine learning models involves several steps that, if not performed carefully, may introduce sources of biases.

For example, many papers use a two-part study. The first part refers to creating the model (e.g., selecting which features best predict the outcome). The second part refers to its validation (e.g., applying it to estimate how well the model works). However, if researchers use the same dataset to create the model and validate it, the estimated accuracy will be inflated (81). Therefore, using a different dataset on model creation (including any feature selection) and validation is mandatory to avoid this bias.

Another example. Due to the difficulty of obtaining large datasets, collaborations between different sites are common. However, machine learnings may “fraudulently” use the differences between sites to predict the outcome. Therefore, these potential effects of the site must be very carefully controlled. Ignoring them may yield an inflated accuracy, even when the models do not really predict (76, 82).

Clinical Utility

A common concern of clinical practitioners is the dubious clinical utility of some machine learning studies (83, 84). On the one hand, machine learning in medical data has proven to be an impressive tool in replicating and automating human processes, such as computer-automated detection (CAD) of lesions on brain scans, body scans, or mammograms (85). However, on the other hand, studies such as detecting whether an MR image belongs to a patient or a healthy control may seem clinically useless (68). We

knowledge that these studies are indeed valuable as a proof-of-concept. However, we should progressively ensure that clinicians find them helpful, i.e., that the question answered by the model aligns well with clinical needs.

In this regard, it is essential to keep a distinction between what is a “model” and what is a “tool.” A model may be necessary for further investigation or methodological purposes. Conversely, a tool should be helpful, feasible, and safe for clinical decision-making in real-world settings (84).

CHALLENGES AND LATEST ADVANCEMENTS

Longitudinal Studies

Methods based on single time-point data can be helpful. Still, changes over time may provide relevant information to create models of what may happen (e.g., if the patient will respond to treatment or have a complication). For example, it is known that patients with a first episode of psychoses show a decrease over time in cortical gray matter when compared with healthy controls (86), or that progressive gray matter volume reduction in the superior temporal gyrus is associated to low improvement in positive psychosis symptoms (87). Having a dataset collected from hundreds or thousands of people with similar conditions over an extended period will enable more complex patterns to be found. These patterns will allow a better future outcome prediction. Therefore, longitudinal studies will be crucial in improving the reliability and performance of mental health decision-helper tools.

Larger Datasets

One of the common first steps when preprocessing neuroimaging data is reducing dimensionality, using expert-designed feature selection or feature extraction. This process boosts the performance of algorithms, but it removes information from the input data. Conversely, modern algorithms like deep learning can use minimally preprocessed input data and take advantage of the subtle patterns usually withdrawn during preprocessing (88). However, although already used in some brain abnormality detection tasks, deep learning has not yet been extensively applied to detect early subjects at risk of developing a disorder or a relevant outcome. A critical reason for not using deep learning algorithms is that they require, in general, substantially larger datasets than other machine learning approaches.

Neuroimaging datasets tend to be hard to acquire. Still, emerging consortia, like the ENIGMA consortium,⁴ are already making it possible to conduct analyses on large datasets otherwise impossible to recruit (76). Indeed, a larger multisite sample not only improves the statistical power of the studies and allows the use of deep learning but also enhances the generalizability of the models to new data.

New Algorithms

Algorithms and methods evolve every day, so maybe the best tool to detect subjects at risk is still testing. In this section, we will

⁴<http://enigma.ini.usc.edu/>

only scratch the surface and review some of the most promising methods in machine learning.

Self-Defining Algorithms

There are many possible machine learning algorithms to apply to a concrete question. The problem is which algorithm or hyperparameters are the best for each outcome. A new methodology called AutoML consists of techniques that can automatically select the appropriate model and its associated hyperparameters to optimize the performance and reliability of the resulting predictions (89). Having the algorithm self-define its characteristics can provide a human-agnostic model definition that is not prone to the biases and assumptions tied to each decision the expert makes when defining a model. It has already been tested in identifying digital phenotyping measures that are more relevant for negative symptoms in psychotic disorders successfully (90).

Combining Knowledge From Other Sources

In other domains, like in computer vision, large datasets exist for general purposes, such as ImageNet (91). But in neuroimaging is not so easy to achieve such a giant dataset. Here is where a technique called Transfer Learning appears. This approach can extract insights obtained in large general-purpose datasets and use that information to improve small dataset model creation (82). This technique has already been tested to improve Alzheimer's disease classification (92). Still, to our knowledge, it has not been used in many other domains like risk estimation.

Explainable Artificial Intelligence

Novel algorithms like Deep Learning are usually considered "black boxes" because networks' decisions are not easily interpretable by humans. Explainable Artificial Intelligence (XAI) seeks to provide an easily understandable solution. For example, in highly complex neural networks used for MRI-based classification, it is not easy to know which voxels have been used to classify between groups; XAI would provide a heat map indicating which were the more relevant zones or voxels used in classification, providing insights into how the network works (93). One approach is the layer-wise relevance propagation (LRP), which produces heatmaps of the contribution of each voxel to the final classification outcome at a single-subject level. When tested in Alzheimer's disease, the voxels reported in the heatmap were concordant with zones associated with AD abnormalities in previous literature (74). It has also been applied to multiple sclerosis, where the lesions are distributed across the brain. The individual heatmaps corresponded to the lesions themselves and non-lesion gray and white matter areas such as the thalamus, which are conventional MRI markers (75). In a study where authors used texture feature maps for classifying participants with SZ, MD patients, and HC, LRP showed which zones contributed to the classification of the deep learning algorithm (94). Another interesting approach to determine which regions contribute the most to classification consists of substitute brain regions by healthy ones generated using variational autoencoders and then see how performance changes (95).

Having tools understandable for humans would make it easier for the researchers, clinicians, and general population to believe in them.

Federated Learning

One obstacle in data sharing for creating larger datasets in MRI is the concern about privacy and confidentiality. And another limitation is that despite having large imaging databases, many images have few labels and therefore do not allow the model to learn much. A trained radiologist must inspect the images to annotate the labels, which can be time-consuming. Both problems can be solved using federated learning since it trains algorithms across multiple health care sites. An algorithm is provided to all the centers and is applied locally at each site. Once the algorithm extracts the information, this knowledge is put together. Using this approach, no private data is shared, and centers can help in the process even if their labeled database is small. Federated learning is a promising technique that upholds patients' privacy and eases the cooperation between health care centers (96).

A Multimodal Approach

Schizophrenia and other mental disorders are known to be caused by a combination of genetic, anatomical, and environmental factors. Therefore, predictions of future outcomes or the early detection of subjects at risk can profit from multimodal approaches, e.g., combining genetic and neuroanatomical factors. Many studies are indeed already using a multimodal approach (97). However, the main problem is that it is still unclear which combination of factors best predict the outcome and how to combine them.

DISCUSSION

This article describes some common techniques and algorithms used in neuroimaging machine learning research. It also reviews some errors and pitfalls that may affect their models' replicability and clinical utility. Finally, it scratches on the surface of some of the topics that can be relevant shortly. One of them may be focusing on acquiring longitudinal data, which may address clinically relevant questions (e.g., whether a patient will respond to one or other treatment). Another topic may be related to emerging promising techniques. For instance, human-agnostic model definition algorithms could provide new assumption-free methods not relying on human choices on algorithm definition. Transfer Learning algorithms could allow using algorithms that have been intensively trained in other fields. Other algorithms may overcome the "black box" machine learning problem. Or, on another note, Federated Learning may ease collaboration between centers, achieving the larger sample sizes required for machine learning.

AUTHOR CONTRIBUTIONS

AS and JR worked equally on the drafting and revision of the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Spanish Ministry of Science, Innovation and Universities/Economy and

Competitiveness/Instituto de Salud Carlos III (CPII19/00009, PI19/00394, and FI20/00047), and co-financed by ERDF Funds from the European Commission (“A Way of Making Europe”).

REFERENCES

- Grover S, Sahoo S. Outcome measurement in schizophrenia: challenges and barriers. In: Shrivastava A, De Sousa A editors. *Schizophrenia Treatment Outcomes*. (Cham: Springer International Publishing) (2020). p. 91–124. doi: 10.1007/978-3-030-19847-3_10
- Laursen TM, Nordentoft M, Mortensen PB. Excess early mortality in schizophrenia. *Ann Rev Clin Psychol*. (2014) 10:425–48. doi: 10.1146/ANNUREV-CLINPSY-032813-153657
- Solmi M, Radua J, Olivola M, Croce E, Soardo L, Salazar de Pablo G, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Mol Psychiatry*. (2021). doi: 10.1038/s41380-021-01161-7
- Howes OD, Whitehurst T, Shatalina E, Townsend L, Onwordi EC, Mak TLA, et al. The clinical significance of duration of untreated psychosis: an umbrella review and random-effects meta-analysis. *World Psychiatry*. (2021) 20:75–95. doi: 10.1002/wps.20822
- Arnone D, McIntosh AM, Tan GMY, Ebmeier KP. Meta-analysis of magnetic resonance imaging studies of the corpus callosum in schizophrenia. *Schizophr Res*. (2008) 101:124–32. doi: 10.1016/j.schres.2008.01.005
- Brugger SP, Howes OD. Heterogeneity and homogeneity of regional brain structure in schizophrenia: a meta-analysis. *JAMA Psychiatry*. (2017) 74:1104–11. doi: 10.1001/JAMAPSYCHIATRY.2017.2663
- Haijma SV, Van Haren N, Cahn W, Koolschijn PCMP, Hulshoff Pol HE, Kahn RS. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. *Schizophr Bull*. (2013) 39:1129–38. doi: 10.1093/schbul/sbs118
- Van Erp TGM, Hibar DP, Rasmussen JM, Glahn DC, Pearlson GD, Andreassen OA, et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol Psychiatry*. (2016) 21:547–53. doi: 10.1038/MP.2015.63
- Vita A, De Peri L, Deste G, Sacchetti E. Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. *Transl Psychiatry*. (2012) 2:e190–. doi: 10.1038/tp.2012.116
- Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. (1995) 20:273–97. doi: 10.1023/A:1022627411411
- Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*. (1970) 12:55–67. doi: 10.1080/00401706.1970.10488634
- Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc*. (1996) 58:267–88. doi: 10.1111/j.2517-6161.1996.tb02080.x
- Ho TK. Random decision forests. In: *Proceedings of the 3rd International Conference on Document Analysis and Recognition*. Montreal (1995). doi: 10.1109/ICDAR.1995.598994
- Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. (2015) 521:436–44. doi: 10.1038/nature14539
- Koutsouleris N, Borgwardt S, Meisenzahl EM, Bottlender R, Möller H-J, Riecher-Rössler A. Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers: results from the FePsy study. *Schizophr Bull*. (2012) 38:1234–46. doi: 10.1093/schbul/sbr145
- Mikolas P, Hlinka J, Koch A, Pitra Z, Frolt T, Spaniel F, et al. Machine learning classification of first-episode schizophrenia spectrum disorders and controls using whole brain white matter fractional anisotropy. *BMC Psychiatry*. (2018) 18:97. doi: 10.1186/S12888-018-1678-Y
- Oh J, Oh BL, Lee KU, Chae JH, Yun K. Identifying schizophrenia using structural MRI with a deep learning algorithm. *Front Psychiatry*. (2020) 11:16. doi: 10.3389/fpsy.2020.00016
- Liu S, Liu S, Cai W, Che H, Pujol S, Kikinis R, et al. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Trans Biomed Eng*. (2015) 62:1132–40. doi: 10.1109/TBME.2014.2372011
- Schnack HG, Nieuwenhuis M, van Haren NEM, Abramovic L, Scheewe TW, Brouwer RM, 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. doi: 10.1016/j.neuroimage.2013.08.053
- Talpaluru A, Bhagwat N, Devenyi GA, Lepage M, Chakravarty MM. Identifying schizophrenia subgroups using clustering and supervised learning. *Schizophr Res*. (2019) 214:51–9. doi: 10.1016/J.SCHRES.2019.05.044
- Lai JW, Ang CKE, Rajendra Acharya U, Cheong KH. Schizophrenia: a survey of artificial intelligence techniques applied to detection and classification. *Int J Environ Res Public Health*. (2021) 18:1–20. doi: 10.3390/ijerph18116099
- Andreou C, Borgwardt S. Structural and functional imaging markers for susceptibility to psychosis. *Mol Psychiatry*. (2020) 25:2773–85. doi: 10.1038/s41380-020-0679-7
- Koutsouleris N, Kambaitz-Illankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, et al. Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: a multimodal, multisite machine learning analysis. *JAMA Psychiatry*. (2018) 75:1156–72. doi: 10.1001/jamapsychiatry.2018.2165
- Rashid B, Calhoun V. Towards a brain-based predictive model of mental illness. *Hum Brain Mapp*. (2020) 41:3468–535. doi: 10.1002/hbm.25013
- Arrasate M, González-Ortega I, García-Alocén A, Alberich S, Zorrilla I, González-Pinto A. Prognostic value of affective symptoms in first-admission psychotic patients. *Int J Mol Sci*. (2016) 17:1039. doi: 10.3390/IJMS17071039
- Hui CLM, Honer WG, Lee EHM, Chang WC, Chan SKW, Chen EYHESM, et al. Predicting first-episode psychosis patients who will never relapse over 10 years. *Psychol Med*. (2019) 49:2206–14. doi: 10.1017/S0033291718003070
- Wunderink L, van Bebbler J, Sytema S, Boonstra N, Meijer RR, Wigman JTW. Negative symptoms predict high relapse rates and both predict less favorable functional outcome in first episode psychosis, independent of treatment strategy. *Schizophr Res*. (2020) 216:192–9. doi: 10.1016/j.schres.2019.12.001
- Bhattacharyya S, Schoeler T, Patel R, di Forti M, Murray RM, McGuire P. Individualized prediction of 2-year risk of relapse as indexed by psychiatric hospitalization following psychosis onset: model development in two first episode samples. *Schizophr Res*. (2021) 228:483–92. doi: 10.1016/j.schres.2020.09.016
- Bowtell M, Eaton S, Thien K, Bardell-Williams M, Downey L, Ratheesh A, et al. Rates and predictors of relapse following discontinuation of antipsychotic medication after a first episode of psychosis. *Schizophr Res*. (2018) 195:231–6. doi: 10.1016/j.schres.2017.10.030
- Bergé D, Mané A, Salgado P, Cortizo R, Garnier C, Gomez L, et al. Predictors of relapse and functioning in first-episode psychosis: a two-year follow-up study. *Psychiatr Serv*. (2016) 67:227–33. doi: 10.1176/appi.ps.201400316
- Schoeler T, Petros N, Di Forti M, Klamerus E, Foglia E, Murray R, et al. Poor medication adherence and risk of relapse associated with continued cannabis use in patients with first-episode psychosis: a prospective analysis. *Lancet Psychiatry*. (2017) 4:627–33. doi: 10.1016/S2215-0366(17)30233-X
- Kopczynska M, Zelek W, Touchard S, Gaughran F, Di Forti M, Mondelli V, et al. Complement system biomarkers in first episode psychosis. *Schizophr Res*. (2019) 204:16–22. doi: 10.1016/j.schres.2017.12.012
- Laskaris L, Zalesky A, Weickert CS, Di Biase MA, Chana G, Baune BT, et al. Investigation of peripheral complement factors across stages of psychosis. *Schizophr Res*. (2019) 204:30–7. doi: 10.1016/j.schres.2018.11.035
- Harrisberger F, Smieskova R, Vogler C, Egli T, Schmidt A, Lenz C, et al. Impact of polygenic schizophrenia-related risk and hippocampal volumes on the onset of psychosis. *Transl Psychiatry*. (2016) 6:e868. doi: 10.1038/tp.2016.143
- Ranlund S, Calafato S, Thygesen JH, Lin K, Cahn W, Crespo-Facorro B, et al. A polygenic risk score analysis of psychosis endophenotypes across brain functional, structural, and cognitive domains. *Am J Med Genet*. (2018) 177:21–34. doi: 10.1002/ajmg.b.32581

36. Vassos E, Di Forti M, Coleman J, Iyegbe C, Prata D, Euesden J, et al. An Examination of polygenic score risk prediction in individuals with first-episode psychosis. *Biol Psychiatry*. (2017) 81:470–7. doi: 10.1016/j.biopsych.2016.06.028
37. Schmidt A, Cappucciati M, Radua J, Rutigliano G, Rocchetti M, Dell'Osso L, et al. Improving prognostic accuracy in subjects at clinical high risk for psychosis: systematic review of predictive models and meta-analytical sequential testing simulation. *Schizophr Bull*. (2017) 43:375–88. doi: 10.1093/schbul/sbw098
38. Cahn W, van Haren NEM, Pol HEH, Schnack HG, Caspers E, Laponder DAJ, et al. Brain volume changes in the first year of illness and 5-year outcome of schizophrenia. *Br J Psychiatry*. (2006) 189:381–2. doi: 10.1192/bjp.bp.105.015701
39. Dazzan P, Reinders AS, Shatzi V, Carletti F, Arango C, Fleischacker W, et al. 31.3 clinical utility of MRI scanning in first episode psychosis. *Schizophr Bull*. (2018) 44:S50–1. doi: 10.1093/schbul/sby014.129
40. Nieuwenhuis M, Schnack HG, van Haren NE, Lappin J, Morgan C, Reinders AA, et al. Multi-center MRI prediction models: predicting sex and illness course in first episode psychosis patients. *Neuroimage*. (2017) 145:246–53. doi: 10.1016/j.neuroimage.2016.07.027
41. Dazzan P, Arango C, Fleischacker W, Galderisi S, Glenthøj B, Leucht S, et al. Magnetic resonance imaging and the prediction of outcome in first-episode schizophrenia: a review of current evidence and directions for future research. *Schizophr Bull*. (2015) 41:574–83. doi: 10.1093/schbul/sbv024
42. Korda AI, Andreou C, Borgwardt S. Pattern classification as decision support tool in antipsychotic treatment algorithms. *Exp Neurol*. (2021) 339:113635. doi: 10.1016/j.expneurol.2021.113635
43. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc*. (2005) 67:301–20. doi: 10.1111/J.1467-9868.2005.00503.X
44. Yassin W, Nakatani H, Zhu Y, Kojima M, Owada K, Kuwabara H, et al. Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl Psychiatry*. (2020) 10:1–11. doi: 10.1038/s41398-020-00965-5
45. Wu Y, Ren P, Chen R, Xu H, Xu J, Zeng L, et al. Detection of functional and structural brain alterations in female schizophrenia using elastic net logistic regression. *Brain Imaging Behav*. (2022) 16:281–90. doi: 10.1007/s11682-021-00501-z
46. Salvador R, Radua J, Canales-Rodríguez EJ, Solanes A, Sarró S, Goikolea JM, et al. Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One*. (2017) 12:e0175683. doi: 10.1371/journal.pone.0175683
47. Zarogianni E, Moorhead TWJ, Lawrie SM. Towards the identification of imaging biomarkers in schizophrenia, using multivariate pattern classification at a single-subject level. *Neuroimage Clin*. (2013) 3:279. doi: 10.1016/J.NICL.2013.09.003
48. Brent BK, Thermenos HW, Keshavan MS, Seidman LJ. Gray matter alterations in schizophrenia high-risk youth and early-onset schizophrenia: a review of structural MRI findings. *Child Adolesc. Psychiatr. Clin. N. Am.* (2013) 22:689–714. doi: 10.1016/j.chc.2013.06.003
49. Zarogianni E, Storkey AJ, Johnstone EC, Owens DGC, Lawrie SM. Improved individualized prediction of schizophrenia in subjects at familial high risk, based on neuroanatomical data, schizotypal and neurocognitive features. *Schizophr Res*. (2017) 181:6–12. doi: 10.1016/j.schres.2016.08.027
50. Zhu F, Liu Y, Liu F, Yang R, Li H, Chen J, et al. Functional asymmetry of thalamocortical networks in subjects at ultra-high risk for psychosis and first-episode schizophrenia. *Eur Neuropsychopharmacol*. (2019) 29:519–28. doi: 10.1016/j.euroneuro.2019.02.006
51. Lalouis PA, Wood SJ, Schmaal L, Chisholm K, Griffiths SL, Reniers RLEP, et al. Heterogeneity and classification of recent onset psychosis and depression: a multimodal machine learning approach. *Schizophr Bull*. (2021) 47:1130–40. doi: 10.1093/SCHBUL/SBAA185
52. Gould IC, Shepherd AM, Laurens KR, Cairns MJ, Carr VJ, Green MJ. Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: a support vector machine learning approach. *Neuroimage*. (2014) 6:229–36. doi: 10.1016/j.nicl.2014.09.009
53. Wenzel J, Haas SS, Dwyer DB, Ruef A, Oeztuerk OF, Antonucci LA, et al. Cognitive subtypes in recent onset psychosis: distinct neurobiological fingerprints? *Neuropsychopharmacology*. (2021) 46:1475–83. doi: 10.1038/s41386-021-00963-1
54. Guo Y, Qiu J, Lu W. Support vector machine-based schizophrenia classification using morphological information from amygdaloid and hippocampal subregions. *Brain Sci*. (2020) 10:E562. doi: 10.3390/brainsci10080562
55. 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. doi: 10.3389/fnhum.2013.00520
56. Wang J, Ke P, Zang J, Wu F, Wu K. Discriminative analysis of schizophrenia patients using topological properties of structural and functional brain networks: a multimodal magnetic resonance imaging study. *Front Neurosci*. (2021) 15:785595. doi: 10.3389/fnins.2021.785595
57. Steardo L, Carbone EA, de Filippis R, Pisanu C, Segura-García C, Squassina A, et al. Application of support vector machine on fmri data as biomarkers in schizophrenia diagnosis: a systematic review. *Front Psychiatry*. (2020) 11:588. doi: 10.3389/fpsy.2020.00588
58. Du Y, Fu Z, Calhoun VD. Classification and prediction of brain disorders using functional connectivity: promising but challenging. *Front Neurosci*. (2018) 12:525. doi: 10.3389/fnins.2018.00525
59. Mika S, Rätsch G, Weston J, Schölkopf B, Müller K-R. Fisher discriminant analysis with kernels. In: *Proceedings of the 1999 9th IEEE Workshop on Neural Networks for Signal Processing (NNSP'99)*. Madison, WI (1999). doi: 10.1016/j.neunet.2007.05.005
60. Karageorgiou E, Schulz SC, Gollub RL, Andreasen NC, Ho B-C, Lauriello J, et al. Neuropsychological testing and structural magnetic resonance imaging as diagnostic biomarkers early in the course of schizophrenia and related psychoses. *Neuroinformatics*. (2011) 9:321–33. doi: 10.1007/s12021-010-9094-6
61. Kasperek T, Thomaz CE, Sato JR, Schwarz D, Janousova E, Marecek R, et al. Maximum-uncertainty linear discrimination analysis of first-episode schizophrenia subjects. *Psychiatry Res*. (2011) 191:174–81. doi: 10.1016/j.psychres.2010.09.016
62. Takayanagi Y, Kawasaki Y, Nakamura K, Takahashi T, Orikabe L, Toyoda E, et al. Differentiation of first-episode schizophrenia patients from healthy controls using ROI-based multiple structural brain variables. *Prog Neuropsychopharmacol Biol Psychiatry*. (2010) 34:10–7. doi: 10.1016/j.pnpbp.2009.09.004
63. Zanetti MV, Schaufelberger MS, Doshi J, Ou Y, Ferreira LK, Menezes PR, et al. Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry*. (2013) 43:116–25. doi: 10.1016/j.pnpbp.2012.12.005
64. Kawasaki Y, Suzuki M, Kherif F, Takahashi T, Zhou S-Y, Nakamura K, et al. Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *Neuroimage Clin*. (2007) 34:235–42. doi: 10.1016/j.neuroimage.2006.08.018
65. Nakamura K, Kawasaki Y, Suzuki M, Hagino H, Kurokawa K, Takahashi T, et al. Multiple structural brain measures obtained by three-dimensional magnetic resonance imaging to distinguish between schizophrenia patients and normal subjects. *Schizophr Bull*. (2004) 30:393–404. doi: 10.1093/oxfordjournals.schbul.a007087
66. Ota M, Sato N, Ishikawa M, Hori H, Sasayama D, Hattori K, et al. Discrimination of female schizophrenia patients from healthy women using multiple structural brain measures obtained with voxel-based morphometry. *Psychiatry Clin Neurosci*. (2012) 66:611–7. doi: 10.1111/j.1440-1819.2012.02397.x
67. Santos PE, Thomaz CE, dos Santos D, Freire R, Sato JR, Louzã M, et al. Exploring the knowledge contained in neuroimages: statistical discriminant analysis and automatic segmentation of the most significant changes. *Artif Intell Med*. (2010) 49:105–15. doi: 10.1016/j.artmed.2010.03.003
68. Winterburn JL, Voineskos AN, Devenyi GA, Plitman E, de la Fuente-Sandoval C, Bhagwat N, et al. Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance imaging and machine learning? A multi-method and multi-dataset study. *Schizophr Res*. (2019) 214:3–10. doi: 10.1016/J.SCHRES.2017.11.038
69. Greenstein D, Malley JD, Weisinger B, Clasen L, Gogtay N. Using Multivariate Machine Learning Methods and Structural MRI to Classify Childhood Onset

- Schizophrenia and Healthy Controls. *Front Psychiatry*. (2012) 3:53. doi: 10.3389/fpsy.2012.00053
70. Schwarz E, Doan NT, Pergola G, Westlye LT, Kaufmann T, Wolfers T, et al. Reproducible grey matter patterns index a multivariate, global alteration of brain structure in schizophrenia and bipolar disorder. *Transl Psychiatry*. (2019) 9:1–13. doi: 10.1038/s41398-018-0225-4
 71. Shahab S, Mulsant BH, Levesque ML, Calarco N, Nazeri A, Wheeler AL, et al. Brain structure, cognition, and brain age in schizophrenia, bipolar disorder, and healthy controls. *Neuropsychopharmacology*. (2019) 44:898–906. doi: 10.1038/s41386-018-0298-z
 72. Ullah F, Ansari SU, Hanif M, Ayari MA, Chowdhury MEH, Khandakar AA, et al. Brain MR image enhancement for tumor segmentation using 3D U-Net. *Sensors*. (2021) 21:7528. doi: 10.3390/s21227528
 73. Shoeibi A, Khodatars M, Jafari M, Moridian P, Rezaei M, Alizadehsani R, et al. Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: a review. *Comput Biol Med*. (2021) 136:104697. doi: 10.1016/j.compbiomed.2021.104697
 74. Böhle M, Eitel F, Weygandt M, Ritter K. Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification. *Front Aging Neurosci*. (2019) 10:194. doi: 10.3389/fnagi.2019.00194/BIBTEX
 75. Eitel F, Soehler E, Bellmann-Strobl J, Brandt AU, Ruprecht K, Giess RM, et al. Uncovering convolutional neural network decisions for diagnosing multiple sclerosis on conventional MRI using layer-wise relevance propagation. *Neuroimage Clin*. (2019) 24:102003. doi: 10.1016/j.nicl.2019.102003
 76. Radua J, Vieta E, Shinohara R, Kochunov P, Quidé Y, Green MJ, et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage*. (2020) 218:116956. doi: 10.1016/j.neuroimage.2020.116956
 77. Sweeney EM, Vogelstein JT, Cuzzocreo JL, Calabresi PA, Reich DS, Crainiceanu CM, et al. A comparison of supervised machine learning algorithms and feature vectors for MS lesion segmentation using multimodal structural MRI. *PLoS One*. (2014) 9:e95753. doi: 10.1371/journal.pone.0095753
 78. Schnack HG, Kahn RS. Detecting neuroimaging biomarkers for psychiatric disorders: sample size matters. *Front Psychiatry*. (2016) 7:50. doi: 10.3389/fpsy.2016.00050/PDF
 79. Radua J, Carvalho AF. Route map for machine learning in psychiatry: absence of bias, reproducibility, and utility. *Eur Neuropsychopharmacol*. (2021) 50:115–7. doi: 10.1016/j.euroneuro.2021.05.006
 80. Rosen M, Betz LT, Schultze-Lutter F, Chisholm K, Haidl TK, Kambeitz-Ilankovic L, et al. Towards clinical application of prediction models for transition to psychosis: a systematic review and external validation study in the PRONIA sample. *Neurosci Biobehav Rev*. (2021) 125:478–92. doi: 10.1016/j.neubiorev.2021.02.032
 81. Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. (2009) 12:535–40. doi: 10.1038/nn.2303
 82. Solanes A, Palau P, Fortea L, Salvador R, González-Navarro L, Llach CD, et al. Biased accuracy in multisite machine-learning studies due to incomplete removal of the effects of the site. *Psychiatry Res Neuroimaging*. (2021) 314:111313. doi: 10.1016/j.psychres.2021.111313
 83. Kim YK, Na KS. Application of machine learning classification for structural brain MRI in mood disorders: critical review from a clinical perspective. *Prog Neuropsychopharmacol Biol Psychiatry*. (2018) 80:71–80. doi: 10.1016/j.pnpbp.2017.06.024
 84. Mechelli A, Vieira S. From models to tools: clinical translation of machine learning studies in psychosis. *NPJ Schizophr*. (2020) 6:4. doi: 10.1038/s41537-020-0094-8
 85. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys*. (2019) 29:102–27. doi: 10.1016/j.zemedi.2018.11.002
 86. Gallardo-Ruiz R, Crespo-Facorro B, Setién-Suero E, Tordesillas-Gutierrez D. Long-term grey matter changes in first episode psychosis: a systematic review. *Psychiatry Investig*. (2019) 16:336. doi: 10.30773/PI.2019.02.10.1
 87. Li R-R, Lyu H-L, Liu F, Lian N, Wu R-R, Zhao J-P, et al. Altered functional connectivity strength and its correlations with cognitive function in subjects with ultra-high risk for psychosis at rest. *CNS Neurosci Ther*. (2018) 24:1140–8. doi: 10.1111/cns.12865
 88. Abrol A, Fu Z, Salman M, Silva R, Du Y, Plis S, et al. Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nat Commun*. (2021) 12:1–17. doi: 10.1038/s41467-020-20655-6
 89. Dafflon J, Pinaya WHL, Turkheimer F, Cole JH, Leech R, Harris MA, et al. An automated machine learning approach to predict brain age from cortical anatomical measures. *Hum Brain Mapp*. (2020) 41:3555–66. doi: 10.1002/hbm.25028
 90. Narkhede SM, Luther L, Raugh IM, Knippenberg AR, Esfahlani FZ, Sayama H, et al. Machine learning identifies digital phenotyping measures most relevant to negative symptoms in psychotic disorders: implications for clinical trials. *Schizophr Bull*. (2021) 48:425–36. doi: 10.1093/schbul/sbab134
 91. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*. (Piscataway, NJ: IEEE) (2009). p. 248–55. doi: 10.1109/CVPR.2009.5206848
 92. Hon M, Khan NM. Towards Alzheimer's disease classification through transfer learning. In: *Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine*. (Dhaka: BIBM) (2017). doi: 10.1109/BIBM.2017.8217822
 93. Fellous JM, Sapiro G, Rossi A, Mayberg H, Ferrante M. Explainable artificial intelligence for neuroscience: behavioral neurostimulation. *Front Neurosci*. (2019) 13:1346. doi: 10.3389/fnins.2019.01346
 94. Korda AI, Ruef A, Neufang S, Davatzikos C, Borgwardt S, Meisenzahl EM, et al. Identification of voxel-based texture abnormalities as new biomarkers for schizophrenia and major depressive patients using layer-wise relevance propagation on deep learning decisions. *Psychiatry Res Neuroimaging*. (2021) 313:111303. doi: 10.1016/j.PSCYCHRESNS.2021.111303
 95. Uzunova H, Ehrhardt J, Kepp T, Handels H. Interpretable explanations of black box classifiers applied on medical images by meaningful perturbations using variational autoencoders. In: Angelini ED, Landman BA editors. *Medical Imaging 2019: Image Processing*. (Bellingham, WA: SPIE) (2019). p. 36. doi: 10.1117/12.2511964
 96. Ng D, Lan X, Yao MM-S, Chan WP, Feng M. Federated learning: a collaborative effort to achieve better medical imaging models for individual sites that have small labelled datasets. *Quant Imaging Med Surg*. (2021) 11:852–7. doi: 10.21037/qims-20-595
 97. Ulaş A, Castellani U, Mirtuono P, Bicego M, Murino V, Cerruti S, et al. Multimodal schizophrenia detection by multiclassification analysis. *Lect Notes Comput Sci*. (2011) 7042:491–8. doi: 10.1007/978-3-642-25085-9_58

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Solanes and Radua. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.