CanJPsychiatry 2013;58(9):499–508

In Review

Neuroimaging-Based Biomarkers in Psychiatry: Clinical Opportunities of a Paradigm Shift

Cynthia H Y Fu, MD, PhD[1]; Sergi G Costafreda, MD, PhD[2]

[1]Reader in Neuroimaging and Affective Disorders, Institute of Psychiatry, King's College London, De Crespigny Park, London, England. Correspondence: Institute of Psychiatry, P074, De Crespigny Park, London SE5 8AF, United Kingdom; cynthia.fu@kcl.ac.uk.

[2]National Institute for Health Research Clinical Lecturer in Old Age Psychiatry, Institute of Psychiatry, King's College London, De Crespigny Park, London, England.

Key Words: biomarkers, neuroimaging, machine learning, personalized, Alzheimer, depression, schizophrenia, magnetic resonance imaging

Received, revised, and accepted January 2013.

Neuroimaging research has substantiated the functional and structural abnormalities underlying psychiatric disorders but has, thus far, failed to have a significant impact on clinical practice. Recently, neuroimaging-based diagnoses and clinical predictions derived from machine learning analysis have shown significant potential for clinical translation. This review introduces the key concepts of this approach, including how the multivariate integration of patterns of brain abnormalities is a crucial component. We survey recent findings that have potential application for diagnosis, in particular early and differential diagnoses in Alzheimer disease and schizophrenia, and the prediction of clinical response to treatment in depression. We discuss the specific clinical opportunities and the challenges for developing biomarkers for psychiatry in the absence of a diagnostic gold standard. We propose that longitudinal outcomes, such as early diagnosis and prediction of treatment response, offer definite opportunities for progress. We propose that efforts should be directed toward clinically challenging predictions in which neuroimaging may have added value, compared with the existing standard assessment. We conclude that diagnostic and prognostic biomarkers will be developed through the joint application of expert psychiatric knowledge in addition to advanced methods of analysis.

Biomarqueurs de neuroimagerie en psychiatrie : possibilités cliniques d'un changement de paradigme

La recherche en neuroimagerie a fourni la preuve des anomalies fonctionnelles et structurelles sous-jacentes des troubles psychiatriques, mais jusqu'ici, elle n'a pas réussi à avoir un impact significatif sur la pratique clinique. Récemment, les diagnostics basés sur la neuroimagerie et les prédictions cliniques tirées d'une analyse d'apprentissage automatique ont démontré un potentiel significatif de traduction clinique. Cette revue présente les concepts clés de cette approche, notamment à quel point l'intégration multivariée des modèles d'anomalies cérébrales est un composant essentiel. Nous passons en revue les résultats récents qui ont des applications potentielles au diagnostic, en particulier, pour les diagnostics précoces et différentiels de la maladie d'Alzheimer et de la schizophrénie, et la prédiction de la réponse clinique au traitement de la dépression. Nous discutons des possibilités cliniques spécifiques et des défis de développement de biomarqueurs pour la psychiatrie en l'absence de standard de référence diagnostique. Nous proposons que les résultats longitudinaux, comme le diagnostic précoce et la prédiction de la réponse au traitement, offrent des possibilités définitives de progrès. Nous proposons que des efforts soient dirigés vers des prédictions cliniquement difficiles dans lesquelles la neuroimagerie peut avoir une valeur ajoutée, comparée à l'évaluation standard existante. Nous concluons que les biomarqueurs diagnostiques et pronostiques seront développés par l'application conjointe du savoir psychiatrique expert et de méthodes d'analyse avancées.

A biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."[1, p 91] Biomarkers are key components of modern medicine; for example, blood glucose monitoring is the cornerstone for the diagnosis and clinical management of diabetes. Nevertheless, decisions for psychiatric disorders are almost entirely based on inferences on self-reported information and the observation of behaviour. Therefore, clinical decisions rely on a patient's ability to collaborate as well as the expertise and experience of the clinician. The resulting potential for ambiguity and bias can result in low reliability for psychiatric diagnoses.[2] Moreover, a diagnosis based on clinical description may not map with etiology and prognosis,[3] which may, in turn, impact negatively on treatment decisions.

It is no longer disputable that psychiatric disorders are brain disorders associated with abnormalities in distributed networks.[4] Neuroimaging provides in vivo access to these abnormal brain circuits, and systematic differences have been observed between healthy people and patients with psychiatric disorders in brain structure, function, and neurochemistry. Brain imaging has substantially increased our knowledge of the pathophysiology of mental illnesses and is a candidate for the development of clinical diagnostic biomarkers. In particular, the potential of neuroimaging-based biomarkers lies not only in diagnosis but also for prognosis. Though clinical features may provide an indication of how well a patient's illness may respond to a particular treatment, biomarkers for predicting clinical response are not currently used in day-to-day practice. A prognostic biomarker could point toward the initiation of more intensive or combined therapies at an earlier stage in patients who have an illness that has been identified as being more difficult to treat, thus reducing the morbidity associated with potentially multiple, poorly effective treatments.

The Development of Neuroimaging-Based Biomarkers

Our review focuses on the key markers needed in clinical applications: diagnosis and prognosis. We review the research studies that have demonstrated neuroimaging-based markers that have the potential for clinical translation. These neuroimaging studies can be broadly grouped into measures of brain structure or brain function. Structural measures include regional brain volumes and measures of white matter, and functional neuroimaging data range from resting-state measures to task-related activation studies. Both structural and functional neuroimaging data can be derived from the brain, as a whole or from specified individual regions. Which neuroimaging measure, combination of neuroimaging measures, or even combination of neuroimaging and clinical measures will provide the best diagnostic and prognostic biomarkers remains an open empirical question. Our review also examines how biomarkers for early and differential diagnostics can be developed, as well as markers that go beyond the usual categorical yes-or-no decision but can offer the probability associated with a particular prediction. We contend that it is by focusing on diagnosis and prognosis that neuroimaging-based biomarkers are more likely to offer added value, compared with the existing clinical assessment.

Abbreviations

AD	antidepressant
BD	bipolar disorder
MCI	mild cognitive impairment
MDD	major depressive disorder
MRI	magnetic resonance imaging
SVM	support vector machine

Highlights

- The development of neurobiological diagnostic markers will require an iterative process, while pursuing higher levels of sensitivity and specificity to perfectly replicate current criteria-based diagnostic categories would be misguided.
- There is potential significant value in neuroimaging-based prediction of clinical response to improve clinical outcome.
- Expert knowledge is crucial for the successful translation of these potential biomarkers to clinical practice.

Limitations

- There are no gold standard diagnostic markers in psychiatric disorders. Consequently, it would be misguided to aim for the highest sensitivity and specificity based on current diagnostic criteria.

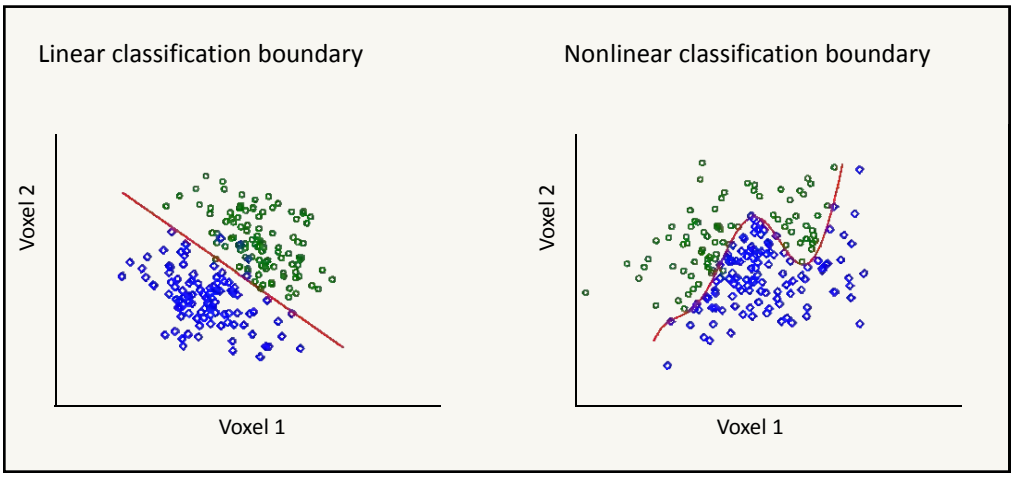
We also review the specific challenges of validating new neuroimaging-based biomarkers for psychiatric conditions as current diagnostic criteria suffer from limited reliability and may be dissociated from neurobiological abnormalities. We propose that the joint application of expert psychiatric knowledge and advanced analytic methods is essential for the development of these much-needed clinical tools. Finally, we discuss the next steps required to maximize the impact of this approach for research and clinical practice.

From Neuroimaging Findings to Clinical Applications

The conventional output from a neuroimaging study is a brain map, which summarizes measures from many subjects. In psychiatry, a typical study reports the significant differences in regional brain responses or structure between a given sample of patients and healthy control subjects. These group-level findings aim to identify the functional and structural brain phenotypes associated with a given disorder. To develop clinical tools, the investigation requires a shift away from considering differences at the group level, such as between patients and control subjects. Instead, the question becomes whether the brain map of a given person is expressing the neural phenotype for a specific disorder or clinical outcome. In other words, the aim is to use brain measurements for the prediction at the individual level regarding diagnosis—whether a given subject is ill or healthy—and prognosis—whether a given patient will show a good or poor clinical response to a particular treatment.

However, a fundamental difficulty is the inherent complexity of brain images that are 3-dimensional arrays of data constituted by a high number of localized measurements (known as its volumetric elements or voxels, for short). Genetics, functional genomics, and proteomics similarly generate detailed, highly dimensional measurements that contain clinically valuable predictive information. All of these fields share similar analytical challenges,[5] as classical statistical methods generally require more observations (subjects) than input variables (for example, voxels or

Figure 1 Improving classification with multivariate methods



This hypothetical classification problem illustrates how multivariate integration can improve accuracy. This example aims to separate 2 groups of subjects (green circles representing depressed patients, blue diamonds as healthy control subjects) based on 2 quantitative neuroimaging measurements (anterior cingulate and amygdala activation). In practice, multivariate integration generally involves more than 2 variables, but the principles illustrated here still apply. None of the 2 features can, in isolation, accurately separate the 2 groups. However, the integration of both measurements using a multivariate linear classification boundary achieves high separation.

genes). Instead, neuroimaging studies typically involve samples sizes in the order of tens to hundreds of subjects consisting of thousands to millions of voxels per subject.

Machine learning is a collection of methods derived from artificial intelligence and statistical learning. These methods have been a key facilitator in developing biomarkers from brain-imaging data as they have proven effective at tackling the analytic challenges of high-dimensional data.[5–7] Such methods have already led to practical commercial applications, for example, face and speech recognition[8] and clinical prediction based on genomic information.[5] In neuroimaging, machine learning methods underpin recent advances in so-called mind-reading, whereby thoughts and intentions are identified from the pattern of brain activation alone.[9]

An important feature of machine learning algorithms is that they are designed to deal with multivariate inputs; in other words, they treat the brain images as patterns rather than considering each voxel in isolation as in conventional analysis methods. Multivariate approaches allow the integration over the whole brain of localized differences, which individually may be too small in magnitude or too variable to reliably separate groups of subjects (Figure 1). As neuroimaging studies in psychiatric disorders tend to reveal abnormalities that affect a network of regions rather than isolated, localized changes,[10] the multivariate approach is able to integrate such patterns of differences leading to improvements in predictive accuracy.

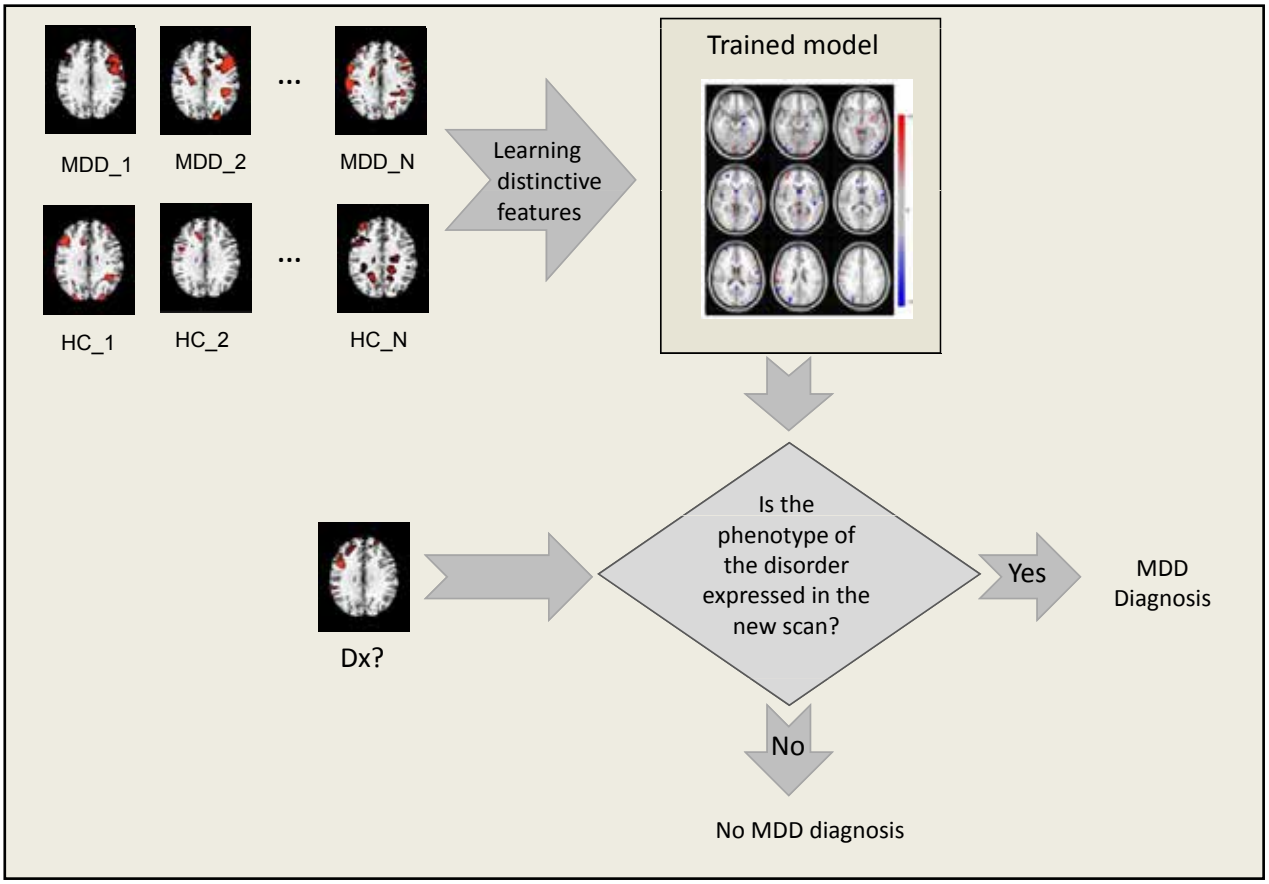
The steps in the analysis involve teaching the machine learning algorithms to recognize samples of brain images, such as patients and control subjects (Figure 2). The algorithm learns the characteristics in these patterns associated with either group. In other words, the algorithm learns the phenotype associated with the characteristic to be predicted, be it diagnostic or prognostic. Once the algorithm has been trained, this knowledge can be applied to novel scans, resulting in subject-specific clinical predictions. For instance, the novel brain image is then classified as belonging to one group or the other, such as patients or control subjects. Further technical details have been recently reviewed.[11,12] Moreover, it is possible to combine different types of data, such as neuroimaging with clinical measures, which may improve the accuracy of the prediction.

Diagnosis, Early Diagnosis, and Differential Diagnosis

The strongest potential for clinical application to date has been in Alzheimer disease, owing to its well-characterized brain atrophy, which begins in the medial temporal lobe and spreads to neocortical regions.[13–15] This atrophic phenotype of Alzheimer disease has been replicated by machine learning analyses.[16–18] Using cases of Alzheimer disease that have been verified by postmortem investigation, high diagnostic accuracy of individual patients has been achieved, with a sensitivity of 95% and specificity of 95%, while expert neuroradiologists achieved only a sensitivity of 88% and specificity of 90% for the same dataset.[19,20] Measures of white matter with diffusion tensor imaging have reported a diagnostic accuracy of 98%.[21]

The clinical diagnosis of Alzheimer disease is made when marked and progressive cognitive impairments are already evident. A diagnosis, though, would be more useful at an earlier stage, while the patient is experiencing few cognitive deficits, such as in the form of self-reported memory loss.

Figure 2 Training and testing of classification models



The phenotype associated with an outcome of interest (in this example, subjects diagnosed with MDD and healthy control subjects) is first learned using the scans of well-characterized people. The trained model is then used to generate predictions based on the scans of new, undiagnosed subjects. The trained model is usually represented by a map depicting the regional abnormalities associated with the prediction most important for classification.

However, currently it is impossible to predict from clinical and neuropsychological testing whether an individual with mild cognitive deficits will subsequently progress toward dementia or will remain stable.[22] Analysis of brain images of people with MCI has the potential to aid with this diagnostic transition.[23–25] Indeed, the development of Alzheimer disease was predicted at an accuracy of 80% from the pattern of changes in hippocampal morphology (n = 30 MCI subjects: sensitivity 80%, specificity 80%),[23] which was validated in a large-scale, multicentre trial at the same level of accuracy (n = 103 MCI subjects: sensitivity 77%, specificity 80%).[24] Importantly, MCI subjects who developed dementia could not have been distinguished at baseline based on clinical, demographic, and neuropsychological features, providing evidence that the neuroimaging-based prediction could add value to the standard diagnostic assessment.

The differential diagnosis of Alzheimer disease from other degenerative dementias also presents as a clinical challenge. Further, the degree of overall brain atrophy may not distinguish between pathologies,[26] instead it may be the pattern of atrophy that is useful for diagnostic classification. In Alzheimer disease, the medial temporal cortex is more specifically affected than in non-Alzheimer pathologies, such as dementia with Lewy bodies and vascular cognitive impairment.[27] A recent 3-way diagnosis has been applied to structural MRI scans of pathologically verified cases of dementia.[28] High accuracy of diagnosis was found for all diagnoses: Alzheimer disease (sensitivity 91%, specificity 84%), Lewy body disease (sensitivity 79%, specificity 99%), and frontotemporal degeneration (sensitivity 84%, specificity 94%).

Functional imaging information has also shown diagnostic potential. Brain activation during language and memory tasks measured by functional MRI distinguishes Alzheimer patients and healthy control subjects with over 80% sensitivity and specificity.[29,30] Classification using resting-state perfusion by single photon emission computed tomography (commonly referred to as SPECT)[31,32] also generated high diagnostic performance (sensitivity 90%, specificity 70%), surpassing the accuracy of a panel of expert neuroradiologists (sensitivity 57%, specificity 82%). Whether functional and structural abnormalities contain distinct diagnostic patterns and can be used for improved clinical performance or whether they express similar and redundant information is an important question to be further explored. Structural MRI has an advantage of already being

routinely used in the assessment of suspected dementia.[33] Therefore, machine learning analysis can be seen as an add-on tool to the usual neuroradiological examination, which has the potential to lead to improved early diagnosis of Alzheimer disease and its differential diagnosis.

MDD is characterized by structural and functional brain abnormalities involving limbic and prefrontal regions.[34,35] Using machine learning analysis, we have shown that the functional correlates of implicitly processing emotional faces is diagnostic for depression (sensitivity 84%, specificity 89%).[36] While the functional neuroanatomy of verbal working memory also shows distinct neural correlates in MDD[37] and generated a statistically significant diagnostic accuracy for depression (sensitivity 65%, specificity 70%),[38] the clinical significance is limited, owing to its low accuracy. Additional sensitivity though may be generated by combining functional imaging tasks of emotional and reward processing.[39] Structural abnormalities are evident in depression, in particular in the hippocampus, which is present in the first episode[40]; however, these provided statistically significant but clinically limited diagnostic potential (sensitivity 65%, specificity 70%) for patients with a moderate severity of depression.[40] The diagnostic accuracy of structural neuroimaging data, though, may be greater in patients who experience a more severe form of the illness.[41]

In schizophrenia, structural MRI has also shown diagnostic potential in identifying patients relative to healthy control subjects, using both whole brain images[42,43] and hippocampal abnormalities.[42–45] However, these studies have generally focused on patients with well-established schizophrenia. Once more, accurate diagnosis of schizophrenia at the early stages is a significant clinical challenge.[46] Characteristic structural brain abnormalities have been identified in people with high levels of schizotypy and prodromal symptoms of psychosis, which have been applied to predict the risk of transition to a psychotic episode.[47,48] As well, the differential diagnosis raises another clinical concern. Distinguishing people with schizophrenia from those with BD and healthy control subjects has been achieved with a sensitivity and specificity of over 90% based on functional abnormalities in prefrontal and temporal cortices and within default neural networks.[49–51]

Neuroimaging-based prediction has also shown promising results in classifying subjects with obsessive–compulsive disorder,[52] autism spectrum disorders,[53] BD,[46,54] and substance abuse.[55] Most of these studies have found abnormalities that were regionally distributed but of relatively small magnitudes, and therefore with much overlap between patients and control subjects for any given region, which machine learning approaches were able to integrate leading to accurate classification.

Predicting Clinical Response to Treatment

Treatment response prediction in depression is a clinical concern in which prognostic biomarkers could have a significant impact. Current treatment decisions are made on a trial-and-error basis, with little, if any, empirically validated guidance for which treatment is likely to be most effective for a particular person. Initial reports of an association of anterior cingulate function with treatment response to ADs[56,57] have been replicated and extended in neuroimaging treatment studies of pharmacotherapy[58] and psychological therapy,[59] with evidence of involvement of areas beyond the anterior cingulate.[60] As well, we have observed distinct predictors of clinical response to pharmacotherapy, compared with psychological treatment. The structural neuroanatomy of depression was highly predictive treatment response to ADs,[40,61] while it was the pattern of functional neural responses to emotional processing that predicted clinical response to cognitive-behavioural therapy.[62] Functional neural correlates of emotional processing or verbal memory showed statistically significant but clinically limited predictive potential for ADs.[36,38] The neural substrate for such predictions was consistent with regions implicated in depression, such as in the amygdala and anterior cingulate.[56–58] These findings suggest that a combination of both structural and functional imaging tests may lead to a useful aid in the clinical management of depression, as well as schizophrenia[63] and other disorders.

Integration With Expert Knowledge I: Defining Clinically Important Targets for Prediction

Expert knowledge is crucial for the successful translation of these potential biomarkers to clinical practice. Clinical expertise is required to formulate useful questions for which neuroimaging could conceivably add value over the existing clinical assessment. For instance, identifying people with established dementia from healthy control subjects does not usually present as a clinical challenge with standard clinical assessments and therefore machine learning is unlikely to add value. However, when doubts exist concerning the etiology of a dementia syndrome, neuroimaging-based prediction can be useful for differential diagnosis.[28] In the early diagnosis of Alzheimer disease, neuroimaging can help to identify the pattern of atrophy characteristic of this disorder, a role already accepted in recently proposed diagnostic criteria. Similarly, while well-established schizophrenia is usually identifiable with a standard clinical assessment, the identification of people at risk who are more likely to develop a full-blown psychotic episode based on their neuroimaging phenotype may offer a real clinical advantage.[48]

A specific difficulty in the development of new biomarkers is the absence of a true diagnostic gold standard for psychiatric disorders. A gold standard should separate subjects with and without the condition with near-perfect accuracy.[64] However, in psychiatry, diagnosis is based on the application of criteria to self-reported information and the observation of behaviour. The reliability of this strategy can be low for many diagnoses.[2] More importantly, current diagnostic classification systems for psychiatry do not explicitly link to etiology, and it is unknown to what extent they capture the underlying biological basis.[3]

In machine learning, successful algorithm training depends on the availability of a sample of subjects with the condition of interest, who are therefore likely to express the neurobiological abnormalities to be learned by the algorithm. Any subjects that have been erroneously classified will introduce unwanted variability in this training process, a blurring of the neurobiological features that will tend to make the resulting test less accurate. Diagnostic misclassification is particularly likely in early diagnosis, before the full clinical picture has appeared. For instance, heterogeneous categories of MCI or prodromal psychosis contain a significant proportion of subjects who do not express the biological markers of the process and will never develop the full condition. More critically, if the aim is to build a more valid diagnostic system rooted in core etiological and pathophysiological realities, then pursuing higher levels of sensitivity and specificity in an attempt to perfectly replicate current criteria-based diagnostic categories would be misguided.[3,65] Further, using a flawed gold standard, such as criteria-based diagnosis to validate a potentially superior biologically-based marker, would yield misleading estimates of its performance.[64]

Instead, a successful training and validation strategy for new biomarkers has to consider the limitations in current diagnostic criteria. A general approach to overcome diagnostic misclassification is to train the model (that is, read the neurobiological abnormalities constituting the phenotype of interest) in subjects for which there is very little diagnostic uncertainty, for instance, because they have remained stable in a particular diagnostic category for some time. For early diagnosis, this involves training the algorithm in patients with an established diagnosis, for example, Alzheimer disease[24] or schizophrenia. While in these subjects the test would have little or no clinical added value, their diagnostic homogeneity should result in a good-quality reading of the neurobiological abnormalities into the trained algorithm. This accurate test can then be applied to subjects for whom diagnostic doubts exist; for instance, because they are suspected to be at the very early stages of a disorder, looking for indications that their neuroimaging phenotype already shows similarities to that of more advanced patients.[24] This approach is particularly suited to predicting clinically important outcomes in which early, prodromal neurobiological changes are suspected to precede the full clinical expression of the disorder. This procedure also has the advantage of conducting the training and testing procedure in entirely separate sets of patients, adding to its validity and the likelihood that the observed predictive performance of the test can be generalized to new patients in whom the condition is suspected.

Another strategy is to use neuroimaging to predict events that are relatively independent of diagnostic criteria but have intrinsic clinical value and can therefore act as external validators of clinical diagnoses.[66,67] One example is the prediction of treatment response. In depression, the demonstration that remission can be predicted based on a brain scan before treatment initiation strongly suggests that neuroimaging is able to capture essential components of its biology.[40,68] As well, the prediction of disease onset, such as using neuroimaging to predict whether someone experiencing memory loss is, in fact, in the early stages of dementia, not only of immediate clinical utility but also offers strong evidence for the validity of the underlying neurobiological abnormalities as diagnostic markers. Additionally, the prediction of prognosis and treatment response is particularly unsatisfactory in many areas of psychiatry, and the usual clinical assessment often does not reveal sufficient information to orient treatment decisions, which are therefore made on a trial-and-error basis. As the value of a new biomarker ultimately lies in their capacity to improve patient outcomes, neuroimaging-based prediction is thus likely to offer significant added value in this area.[69]

Integration With Expert Knowledge II: Defining What to Measure

Nevertheless, the role of neuroimaging-based biomarkers for the diagnosis of psychiatric disorders is not necessarily limited to replicating diagnostic criteria systems. We expect that there will be an iterative process, in which previous neuroimaging-based prediction results can generate new hypotheses and be incorporated as prior knowledge for future studies. Raw brain measurements are not usually employed to train the machine learning algorithms, but rather highly preprocessed data are used instead. In most applications, measurements from many brain regions may not be valuable for the prediction task. This redundant input should be discarded while identifying and focusing the learning process on the most promising features. Because these spurious features have no real discriminative power in the population, using such irrelevant voxels to develop a predictive model will compromise its capacity to generate accurate predictions for new subjects.[6,70,71] If there is a known set of affected regions, the classifier can be focused on measurements from these regions of interest, while ignoring measurements from other regions.[72] In Alzheimer disease, the preprocessing can thus involve segmentation of the brain regions known to suffer atrophy in the earlier stages of the illness, namely, in the hippocampus.[73] Therefore, selection of the appropriate imaging modality and preprocessing strategy is crucial for successful prediction as it reflects previous expert knowledge. In most disorders, though, the best modality, preprocessing strategy, and regions of interest are not fully known a priori. In these situations, prior knowledge can be combined with data-driven selection of the most important input features.[68]

Next Steps

Machine learning applied to neuroimaging is a novel approach that has the capacity to generate biomarkers with clinical applicability. The examples have focused on the results most likely to translate into clinical tools, and they are likely to be only a small subset of the potential of this methodology. Whether a group-level difference in brain patterns translates into statistically significant but more

importantly a clinically relevant individual-level prediction is an empirical, testable hypothesis that should be addressed as a complement to conventional analysis.[24,51] This inclusive approach to biomarker discovery in neuroimaging mirrors the strategies currently used in similar high-dimensional fields, such as genetics and *-omics*.[5]

To realize this translational potential, the potential biomarker needs to be taken from the discovery stage, to the development of standardized clinical assay, and to clinical implementation.[74] The next step after discovery is to confirm the initial results in an external sample; that is, the demonstration that the biomarker generalizes and is also useful in other patients. The diagnostic value of medial temporal atrophy in Alzheimer disease has been well studied and is considered to be a confirmatory diagnostic criterion.[74] In depression, the findings that functional abnormalities have diagnostic potential[36] and the predictive potential of grey matter abnormalities[40] has been confirmed in independent samples.[39,61] Validation studies are essential and should include subjects representative of the population, compared with the so-called clean, comorbidity-free, clear-cut patients who are normally used in proof-of-principle studies.[73] Multicentre designs are particularly convincing, because they can test whether that the prediction is robust to differences in prevalence, recruitment, and clinical management.[20,73]

With standardization of acquisition methods and analysis,[75] the predictive models obtained with machine learning analysis in one study can be stored in a database and tested with different samples. A repository consisting of trained models that have shown predictive potential in one sample can be accessed by other researchers for test purposes and would accelerate progress. Well-documented, user-friendly, open-source machine learning toolboxes have been developed (for example, Princeton Multi-Voxel Pattern Analysis toolbox [MVPA][76] and its Python version [PyMVPA][77]), which can analyze most types of neuroimaging data using various approaches, and offer a relatively pain-free approach for clinical scientists to test the predictive value of their experiments.[9] There are also numerous mathematical methods for pattern recognition that have proved their value when applied to neuroimaging. It remains an area of active inquiry in machine learning research whether specific algorithms work best for specific high-dimensional datasets. The predominant technique has been SVM classification.[7,78] SVM finds an optimal boundary between groups by focusing not on the full sample, but only on those examples that are difficult to classify. This leads to good theoretical generalization properties even in high-dimensional classification problems, which have been empirically effective in different areas of medicine, including for the diagnosis of breast tumours in mammographies,[79] and cancer diagnosis and prediction of response to chemotherapy using gene expression data,[7,80,81] in addition to neuroimaging. While some concerns have been raised that relatively long analysis time (currently in the order of days) may be an impediment to the practical application of neuroimaging-based diagnosis,[12] we would assert that such concerns are overstated as this period is comparable to many other laboratory tests, while psychiatric emergencies can be continued to be treated as per current procedures.

An important step, though, is to go beyond dichotic prediction, such as patient compared with control subject, or responder compared with nonresponder. We have proposed a general probabilistic approach to neuroimaging-based classification, which produces accurate predictions along with a confidence level, the probability that the prediction is actually correct for that specific patient.[68] Developments in pattern regression also allow the prediction of continuous variables,[82] paralleling the dimensional outcomes commonly used in clinical psychiatry, such as rating or symptom scales.

A substantial effort is required to translate the outcome of the machine learning algorithm into useful information for the clinician. Most studies present a figure illustrating which brain areas contained the information on which the prediction is based alongside the usual performance metrics of sensitivity and specificity. This brain map can help convince the clinician of the robustness of the predictions, particularly if the areas highlighted have been independently linked to the outcome by other reports. For instance, our work using structural MRI to predict treatment response in depression confirms the well-replicated link between anterior cingulate and AD response.[40,60] New insights into the mechanisms of brain disorders will also likely be revealed. Machine learning prediction, in common with other multivariate, pattern-based approaches, is also more sensitive to distributed patterns of changes that are too weak to survive the strict multiple comparison correction of usual mass-univariate approaches, as in conventional analysis. Increased sensitivity may account for the more complex patterns of abnormalities that machine learning studies generally reveal relative to conventional group analysis.[36,48,83]

The ethical, social, and clinical arguments for the development of biomarkers for diagnosis and prognosis in psychiatry are clear and imperative. It is foreseeable that the clinical assessment for some psychiatric disorders will soon include a brain imaging scan. In fact, the addition of these new analysis techniques to the usual diagnostic work up for dementia is likely to be the first application of neuroimaging-based prediction implemented in clinical practice.

In summary, neuroimaging research has revealed that psychiatric disorders are associated with complex, distributed, multimodal patterns of brain abnormalities. We have discussed how machine learning can provide a unique link between in vivo brain measurements of individual patients and their symptoms, behaviours, and clinical outcomes. This innovation opens the way for the translation of neuroimaging findings into clinical tools for psychiatry and has the potential to make a significant contribution to psychiatric classification by identifying the neurobiology in an iterative process.

Acknowledgements

Dr Fu acknowledges support from a Higher Education Funding Council of England Clinical Senior Lecturer Award and Medical Research Council (United Kingdom) Discipline Hopping Grant. Dr Costafreda acknowledges support from the National Institute for Health Research Specialist Biomedical Research Centre for Mental Health award to the South London and Maudsley National Health Service Foundation Trust and the Institute of Psychiatry, King's College London.

The Canadian Psychiatric Association proudly supports the In Review series by providing an honorarium to the authors.

References

1. Biomarkers Definitions Working Group NIH. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89–95.
2. Kraemer HC, Kupfer DJ, Clarke DE, et al. DSM-5: how reliable is reliable enough? Am J Psychiatry. 2012;169(1):13–15.
3. Kupfer DJ, First MB, Regier DA, editors. Neuroscience research agenda to guide development of a pathophysiologically based classification system. A research agenda for DSM-V. Washington (DC): American Psychiatric Press; 2002.
4. Insel T, Cuthbert B, Garvey M, et al. Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167(7):748–751.
5. Clarke R, Ressom HW, Wang A, et al. The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. Nat Rev Cancer. 2008;8(1):37–49.
6. Duin RPW. Classifiers in almost empty spaces. 15th International Conference on Pattern Recognition; Barcelona, Spain; September 3–7, 2000. p 1–7.
7. Noble WS. What is a support vector machine? Nat Biotechnol. 2006;24(12):1565–1567.
8. Mitchell TM. The discipline of machine learning. Technical report. Pittsburgh (PA): Carnegie Mellon University, School of Computer Science, Machine Learning Department; 2006. Contract number CMU-ML-06–108.
9. Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage. 2009;45(1 Suppl 1):S199–S209.
10. Costafreda SG. Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. Front Neuroinform. 2009;3:33.
11. Kloppel S, Abdulkadir A, Jack CR Jr, et al. Diagnostic neuroimaging across diseases. Neuroimage. 2012;61(2):457–463.
12. Orrù G, Pettersson-Yeo W, Marquand AF, et al. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. Neurosci Biobehav Rev. 2012;36(4):1140–1152.
13. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 1991;82(4):239–259.
14. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007;6(8):734–746.
15. Frisoni GB, Fox NC, Jack CR, et al. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010;6(2):67–77.
16. Fan Y, Batmanghelich N, Clark CM, et al. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. Neuroimage. 2008;39(4):1731–1743.
17. Gutman B, Wang Y, Morra J, et al. Disease classification with hippocampal shape invariants. Hippocampus. 2009;19(6):572–578.
18. Teipel SJ, Born C, Ewers M, et al. Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment. Neuroimage. 2007;38(1):13–24.
19. Kloppel S, Stonnington CM, Barnes J, et al. Accuracy of dementia diagnosis: a direct comparison between radiologists and a computerized method. Brain. 2008;131(Pt 11):2969–2974.
20. Kloppel S, Stonnington CM, Chu C, et al. Automatic classification of MR scans in Alzheimer's disease. Brain. 2008;131(Pt 3):681–689.
21. Graña M, Termenon M, Savio A, et al. Computer aided diagnosis system for Alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation. Neurosci Lett. 2011;502(3):225–229.
22. Fellows L, Bergman H, Wolfson C, et al. Can clinical data predict progression to dementia in amnestic mild cognitive impairment? Can J Neurol Sci. 2008;35(3):314–322.
23. Ferrarini L, Frisoni GB, Pievani M, et al. Morphological hippocampal markers for automated detection of Alzheimer's disease and mild cognitive impairment converters in magnetic resonance images. J Alzheimers Dis. 2009;17(3):643–659.
24. Costafreda SG, Dinov ID, Tu Z, et al. Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. Neuroimage. 2011;56(1):212–219.
25. Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. Neuroimage. 2009;44(4):1415–1422.
26. O'Brien JT, Paling S, Barber R, et al. Progressive brain atrophy on serial MRI in dementia with Lewy bodies, AD, and vascular dementia. Neurology. 2001;56(10):1386–1388.
27. Burton EJ, Barber R, Mukaetova-Ladinska EB, et al. Medial temporal lobe atrophy on MRI differentiates Alzheimer's disease from dementia with Lewy bodies and vascular cognitive impairment: a prospective study with pathological verification of diagnosis. Brain. 2009;132(Pt 1):195–203.
28. Vemuri P, Simon G, Kantarci K, et al. Antemortem differential diagnosis of dementia pathology using structural MRI: differential-STAND. Neuroimage. 2011;55(2):522–531.
29. Ford J, Farid H, Makedon F, et al. Patient classification of fMRI activation maps. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2003, 6th International Conference, Montreal, QC; November 15–18, 2003; Proceedings. Springer; 2003. p 58–66.
30. Kontos D, Megalooikonomou V, Pokrajac D, et al. Extraction of discriminative functional MRI activation patterns and an application to Alzheimer's disease. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2004, 7th International Conference, Saint-Malo, France; September 26–29, 2004; Proceedings; Springer; 2004. p 727–735.
31. Stoeckel J, Ayache N, Malandain G, et al. Automatic classification of SPECT images of Alzheimer's disease patients and control subjects. Lect Notes Comput Sci. 2004:654–662.
32. Fung G, Stoeckel J. SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. Knowl Inf Syst. 2007;11(2):243–258.
33. National Institute for Health and Clinical Excellence (NICE). NICE guidelines for dementia. London (GB): NICE; 2010.
34. Fitzgerald PB, Laird AR, Maller J, et al. A meta-analytic study of changes in brain activation in depression. Hum Brain Mapp. 2008;29(6):683–695.
35. Koolschijn PC, van Haren NE, Lensvelt-Mulders GJ, et al. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Hum Brain Mapp. 2009;30(11):3719–3735.
36. Fu CHY, Mourao-Miranda J, Costafreda SG, et al. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. Biol Psychiatry. 2008;63(7):656–662.

37. Walsh ND, Williams SCR, Brammer MJ, et al. A longitudinal functional magnetic resonance imaging study of verbal working memory in depression after antidepressant therapy. Biol Psychiatry. 2007;62(11):1236–1243.
38. Marquand AF, Mourao-Miranda J, Brammer MJ, et al. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. Neuroreport. 2008;19(15):1507–1511.
39. Hahn T, Marquand AF, Ehlis A-C, et al. Integrating neurobiological markers of depression. Arch Gen Psychiatry. 2011;68(4):361–368.
40. Costafreda SG, Chu C, Ashburner J, et al. Prognostic and diagnostic potential of the structural neuroanatomy of depression. PLoS One. 2009;4(7):e6353.
41. 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(1):64–71.
42. Golland P, Grimson WEL, Shenton ME, et al. Detection and analysis of statistical differences in anatomical shape. Med Image Anal. 2005;9(1):69–86.
43. Yushkevich P, Joshi S, Pizer SM, et al. Feature selection for shape-based classification of biological objects. Inf Process Med Imaging. 2003;2732:114–125.
44. Ford J, Shen L, Makedon F, et al. A combined structural-functional classification of schizophrenia using hippocampal volume plus fMRI activation. Second Joint EMBS-BMES Conference 2002—24th Annual International Conference of the Engineering in Medicine and Biology Society and Annual Fall Meeting of the Biomedical Engineering Society; October 23-26, 2002, Houston (TX). p 48–49.
45. Shen L. A surface-based approach for classification of 3D neuroanatomic structures. Intell Data Anal. 2004;8(6):519–542.
46. Phillips LJ, Yung AR, Yuen HP, et al. Prediction and prevention of transition to psychosis in young people at incipient risk for schizophrenia. Am J Med Genet. 2002;114(8):929–937.
47. Shinkareva SV, Ombao HC, Sutton BP, et al. Classification of functional brain images with a spatio-temporal dissimilarity map. Neuroimage. 2006;33(1):63–71.
48. Koutsouleris N, Meisenzahl EM, Davatzikos C, et al. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch Gen Psychiatry. 2009;66(7):700.
49. Calhoun VD, Maciejewski PK, Pearlson GD, et al. Temporal lobe and "default" hemodynamic brain modes discriminate between schizophrenia and bipolar disorder. Hum Brain Mapp. 2008;29(11):1265–1275.
50. Costafreda SG, Fu CH, Picchioni M, et al. Increased inferior frontal activation during word generation: a marker of genetic risk for schizophrenia but not bipolar disorder? Hum Brain Mapp. 2009;30(10):3287–3298.
51. Costafreda SG, Fu CHY, Picchioni M, et al. Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder. BMC Psychiatry. 2011;11:18.
52. Soriano-Mas C, Pujol J, Alonso P, et al. Identifying patients with obsessive–compulsive disorder using whole-brain anatomy. Neuroimage. 2007;35(3):1028–1037.
53. Ecker C, Marquand A, Mourão-Miranda J, et al. Describing the brain in autism in five dimensions—magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. J Neurosci. 2010;30(32):10612–10623.
54. Calhoun V, Pearlson G, Kiehl K. Automated classification of bipolar, schizophrenic, and healthy individuals via multiple spatial ICA functional brain modes. Machine Learning for Signal Processing, Proceedings of the 16th IEEE Signal Processing Society Workshop; September 6–8, 2006; Maynooth, Ireland. p 371–376.
55. Zhang L, Samaras D, Tomasi D, et al. Machine Learning for Clinical Diagnosis from Functional Magnetic Resonance Imaging. IEEE Computer Society Conference on Computer Vision and Pattern Recognition; June 20–26, 2005; San Diego, CA. p 1211–1217.
56. Buchsbaum MS, Joseph W, Siegel BV, et al. Effect of sertraline on regional metabolic rate in patients with affective disorder. Biol Psychiatry. 1997;41(1):15–22.
57. Mayberg HS, Brannan SK, Mahurin RK, et al. Cingulate function in depression: a potential predictor of treatment response. Neuroreport. 1997;8(4):1057–1061.
58. Fu CH, Williams SC, Cleare AJ, et al. Attenuation of the neural response to sad faces in major depression by antidepressant treatment: a prospective, event-related functional magnetic resonance imaging study. Arch Gen Psychiatry. 2004;61(9):877–889.
59. Fu CHY, Williams SCR, Cleare AJ, et al. Neural responses to sad facial expressions in major depression following cognitive behavioral therapy. Biol Psychiatry. 2008;64(6):505–512.
60. Fu CH, Steiner H, Costafreda SG. Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. Neurobiol Dis. 2013;52:75–83.
61. Gong Q, Wu Q, Scarpazza C, et al. Prognostic prediction of therapeutic response in depression using high-field MR imaging. Neuroimage. 2011;55:1497–1503.
62. Costafreda SG, Khanna A, Mourao-Miranda J, et al. Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. Neuroreport. 2009;20(7):637–641.
63. Khodayari-Rostamabad A, Hasey GM, MacCrimmon DJ, et al. A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy. Clin Neurophysiol. 2010;121(12):1998–2006.
64. Rutjes A, Reitsma J, Coomarasamy A, et al. Evaluation of diagnostic tests when there is no gold standard. A review of methods. Health Technol Assess. 2007;50(11):iii, ix–51.
65. Regier DA, Narrow WE, Kuhl EA, et al. The conceptual development of DSM-V. Am J Psychiatry. 2009;166(6):645–650.
66. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. Am J Psychiatry. 1970;126(7):983–987.
67. Kendler KS. Toward a scientific psychiatric nosology. Strengths and limitations. Arch Gen Psychiatry. 1990;47(10):969–973.
68. Nouretdinov I, Costafreda SG, Gammerman A, et al. Machine learning classification with confidence: application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression. Neuroimage. 2011;56(2):809–813; Epub 2010 May 17.
69. Ferrante di Ruffano L, Hyde CJ, McCaffery KJ, et al. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. BMJ. 2012;344:e686.
70. Jain A, Duin R, Mao J. Statistical pattern recognition: a review. IEEE Trans Pattern Anal Mach Intell. 2000;22(1):4–37.
71. Higdon R, Foster NL, Koeppe RA, et al. A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG-PET imaging. Stat Med. 2004;23(2):315–326.
72. Huppertz HJ, Kröll-Seger J, Danek A, et al. Automatic striatal volumetry allows for identification of patients with chorea-acanthocytosis at single subject level. J Neural Transm. 2008;115(10):1393–1400.
73. Dubois B, Feldman HH, Jacova C et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007;6(8):734–746.
74. Pepe MS, Feng Z, Janes H, et al. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. J Natl Cancer Inst. 2008;100(20):1432–1438.
75. Henley SM, Ridgway GR, Scahill RI, et al. Pitfalls in the use of voxel-based morphometry as a biomarker: examples from Huntington disease. Am J Neuroradiol. 2010;31(4):711–719.

76. Detre GJ, Polyn SM, Moore CD, et al. The Multi-Voxel Pattern Analysis (MVPA) toolbox. Poster presented at the Annual Meeting of the Organization for Human Brain Mapping. 2006 Florence (IT). Available from: http://www.csbmb.princeton.edu/mvpa.
77. Hanke M, Halchenko YO, Sederberg PB, et al. PyMVPA: a Python toolbox for multivariate pattern analysis of fMRI data. Neuroinformatics. 2009;7(1):37–53. Available from: http://www.pymvpa.org.
78. Vapnik V. The nature of statistical learning theory. New York (NY): Springer-Verlag; 1995.
79. Chang RF, Wu WJ, Moon WK, et al. Support vector machines for diagnosis of breast tumors on US images. Acad Radiol. 2003;10(2):189–197.
80. Lee Y, Mangasarian O, Wolberg W. Breast cancer survival and chemotherapy: a support vector machine analysis. Discrete mathematical problems with medical applications: DIMACS Workshop, December 8–10, 1999 [book]. [place of publication unknown]: American Mathematical Society; 2000. p 1–10.
81. Furey TS, Cristianini N, Duffy N, et al. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics. 2000;16(10):906–914.
82. Wang Y, Fan Y, Bhatt P, et al. High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. Neuroimage. 2010;50(4):1519–1535.
83. Davatzikos C, Shen D, Gur RC, et al. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. Arch Gen Psychiatry. 2005;62(11):1218–1227.