The quest for biomarkers in Schizophrenia: from neuroimaging to machine learning
The quest for biomarkers in Schizophrenia: from neuroimaging to machine learning

Miguel Bajouco1,2, David Mota1,2, Manuel Coroa1,2, Salomé Caldeira2, Vitor Santos1,2, and Nuno Madeira1,2

Special Issue on the Neurobiology of Mental Illness

Abstract

Schizophrenia is a severe mental disorder and one of the leading causes of disease burden worldwide. It represents a source of significant suffering and disability to the affected individuals, and is associated with substantial societal and economical costs.

The diagnosis of schizophrenia still depends exclusively on the detection of symptoms that are also present in other mental disorders. This situation causes overlapping of the boundaries of the diagnostic categories and constitutes a source of diagnostic errors. Moreover, current treatment algorithms do not take into account the substantial interindividual variability in response to antipsychotic drugs. As a result, around one-third of patients are treatment-resistant to first line antipsychotic drugs. This deleterious consequence is associated with poor individual outcomes and elevated healthcare costs.

Neuroimaging research in schizophrenia has shed some light in a vast array of structural and functional connectivity abnormalities and neurochemical (dopamine and glutamate) imbalances, which may constitute ‘organic surrogates’ of this disorder. However, the neuroimaging field, so far, has not been able to identify biomarkers that could facilitate early detection and allow individualised treatment management.

This paper reviews neuroimaging studies from different modalities that may provide relevant biomarkers for schizophrenia. We discuss how the current application of novel Machine Learning methods to the analyses of imaging data is allowing the translation of such findings into potential biomarkers enabling the prediction of clinical outcomes at the individual level, towards the development of innovative and personalised treatment strategies.

Keywords: Schizophrenia, Biomarkers, Neuroimaging, Machine-Learning.
Introduction

Schizophrenia has an estimated lifetime prevalence of 0.30–0.66% in the general population [1], representing a source of substantial distress to affected individuals and their relatives. It has been considered one of the leading factors of global disease burden, reducing the life expectancy of patients by more than 18 years due to factors such as increased substance use, poverty, neglect of personal well-being, smoking, metabolic syndrome, and suicide [2, 3].

Schizophrenia’s definition is still based exclusively on clinical criteria, which have remained virtually identical over the last 100 years, grounding its diagnosis on century-old diagnostic categories [4]. Moreover, clinical presentation is heterogeneous not only when comparing between different individuals but also depending on the illness stage, varying from positive symptoms (hallucinations, delusions, disturbances of self, thought disorders, psychomotor abnormalities) to negative, affective and cognitive symptoms [5]. There is also considerable diversity in outcome: 35 to 65% of patients show an undulating course (i.e., episodic and neither continuous nor episodic) with good outcome and 12% to 32% present a chronic course with poor outcome [6]. Antipsychotic drugs are the mainstay of treatment for psychosis, yet there is considerable heterogeneity in their therapeutic efficacy [7, 8]. Consequently, adequate response to treatment is not universal in schizophrenia. Although the majority of patients respond to typical or atypical non-clozapine antipsychotics, roughly a third of patients do not respond well and are considered treatment-resistant [9]. Nonresponse to first-line antipsychotic agents is associated with poor functional outcomes and poses important challenges to health care systems, including up to a tenfold increase in total health resource utilization [10]. Treatment algorithms for patients in their first psychotic episode are devoid of prognostic measures, and therefore treatments are still chosen by “trial and error”, without reference to, or guidance from, the biological background of the individual. Moreover, when patients do not have an adequate clinical response to treatment, they may endure prolonged periods of untreated illness. Current practice suggests a need for reliable, biologically based prognostic measures of treatment response to antipsychotic agents.

In the last decade, considerable research efforts revealed a vast array of chemical [11] functional [12] and structural brain abnormalities [13] that may constitute the ‘organic surrogate’ of the illness [14]. The field of neuroimaging leveraged unprecedented insight into brain biology, showing promising progress towards the identification of putative biomarkers, which could be used for early diagnosis, treatment planning, and monitoring of disease progression. However, to date, the results of neuroimaging studies have not been successfully integrated into the diagnostic process as disease biomarkers operating at the single-subject level [14, 15]. As a result, psychiatrists are still constrained to rely on traditional and often ineffective diagnostic and prognostic tools. One of the reasons for the limited impact of the findings on clinical practice is that neuroimaging studies have typically reported differences between patients and controls, but at group level; in contrast, doctors working in a psychiatric ward have to make clinical decisions about individuals [16]. To overcome these methodological drawbacks, an increasing number of studies have now applied novel multivariate Machine Learning methods, which provide statistical inferences at the individual level [14, 17]. This progress may pave the way to improved diagnosis and development of innovative and personalised treatment strategies.

In this paper, we review recent neuroimaging findings on structural and functional connectivity as well on the chemistry of the brain of individuals with schizophrenia and discuss how obstacles for translating them into clinical practice are being challenged with advanced methods of data analysis.

The development of techniques that allow clinicians to improve early diagnosis and prediction of outcome and treatment is a major translational goal for research in this field.

Neurochemical biomarkers

Two of the most influential hypotheses concerning the underlying neurobiology of the disorder involve dopamine and glutamate [11]. The first evidence supporting these hypotheses goes back to the 1970s originating from in vitro studies on dopamine receptors and antipsychotics. Since then, molecular imaging studies with both Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) extended these in vitro findings to patients. PET and SPECT imaging allows the in vivo quantification of dopamine synthesis, the degree of dopamine release in response to stimuli, and the availability of the post-synaptic dopaminergic receptors and transporters [18]. Studies with these imaging modalities provided compelling evidence of abnormal dopaminergic function in schizophrenia with increased presynaptic dopamine availability and dopamine release [11, 12]. In addition, research with PET has validated the link between elevated dopaminergic synthesis and specific psychotic symptoms [19] and their severity. Importantly, it has been demonstrated that evaluation of striatal [18F]-DOPA uptake allows for classification of schizophrenia patients and controls with approximately 60–90% sensitivity and specificity, depending on the quantification method used [20].

The dopaminergic model of schizophrenia was also strongly based on early evidence demonstrating that antipsychotic drugs cross the blood–brain barrier and block D2/3 striatal receptors in vivo at clinically effective doses. These data extended, and have provided the foundation for studies in which the relationship between D2 occupancy and clinical response could be established [11]. Accordingly, in the treatment of schizophrenia with antipsychotic drugs,
an average 60% occupancy of brain D2 dopamine receptors is required to produce a therapeutic response [21]. However, approximately, a third of patients with schizophrenia show limited if any response to antipsychotic medications, despite high levels of dopamine D2 receptor blockade, with exception to clozapine. This fraction of non-responders has been labelled "treatment-resistant schizophrenia" [22]. The current evidence supports the notion that dysfunction of glutamatergic neurotransmission could be responsible for treatment-resistance. Indeed, glutamatergic abnormalities are thought to contribute to the pathophysiology of schizophrenia and the prevailing hypothesis is for the primary involvement of N-Methyl-D-Aspartate (NMDA) receptor dysfunction [11]. Although the direct evidence of NMDA receptor hypofunction in schizophrenia is still limited, one SPECT imaging study using an NMDA receptor tracer reported a reduction in NMDA receptor binding in medication-free patients with schizophrenia [23]. In addition, human Proton Magnetic Resonance Spectroscopy ([1H]-MRS) studies have shown that ketamine, a NMDA antagonist that induces a schizophrenia-like phenotype, increases glutamate (or glutamine) levels in human prefrontal cortex [24, 25]. The mechanism that explains these changes is still controversial, although it has been shown that inhibition of NMDA receptors leads to reduced GABAergic interneuron function, resulting in disinhibition of pyramidal neurons and excessive glutamate spillover [26]. [1H]-MRS studies at field strengths of 3-Tesla or above in antipsychotic-naive or minimally medicated first-episode psychosis patients have identified increases in the glutamate metabolite glutamine (Gln) or the Gln/glutamate ratio in the anterior cingulate cortex [27, 28], increases in Gln in the thalamus [28], and increases in glutamate in the associative striatum [29]. This suggests that brain glutamatergic activity is increased in the early stages of the disorder, when patients are usually experiencing acute psychotic symptoms [30].

Glutamatergic levels traced by [1H]-MRS and dopamine synthesis capacity measured by [18F]-DOPA PET study have found that patients treated with clozapine (treatment-resistant schizophrenia) show lower dopamine synthesis capacity than patients who have responded to first-line treatment [34]. Based on the findings described above it has been argued that elevated anterior cingulate glutamate may constitute a stable neurobiological trait in treatment-resistant patients, which is preserved after a number of treatment courses with different non-clozapine medications and that the effectiveness of clozapine or novel glutamatergic antipsychotics in treatment-resistant patients might be related to attenuation of glutamatergic dysfunction [33]. To address this question, future studies should use a prospective design, measuring glutamate in medication-naive patients and repeating it when treatment resistance has been determined. Notwithstanding, it has been hypothesized that "treatment-resistant" individuals have a different strain of schizophrenia not characterized by mesolimbic dopaminergic dysfunction. Accordingly, Howes and Kapur [35] proposed two types of schizophrenia: type A (hyperdopaminergic) which is characterized by elevated striatal dopamine synthesis and release capacity, creating a state of hyperdopaminergia which underlies the onset of the disorder, increases as the illness first develops and leads to the development of symptoms; and type B (normodopaminergic) where dopaminergic function is normal and symptoms are not related to dopaminergic function. It is expected that another neurochemical dysfunction underlies type B schizophrenia, being glutamatergic function the most likely.

In summary, we might say that neuroimaging studies suggest that a different neurobiology may underlie treatment-responsive and treatment-resistant schizophrenia. If elevated glutamate levels are found to be a stable neurobiological trait of treatment resistance, they could be used to fast-track such patients to clozapine and/or towards newer glutamatergic compounds under current development.

**Connectivity biomarkers**

Schizophrenia may be considered a disorder of brain connectivity. In fact, the name itself implies a splitting (schizeh) of the mind’s (phren) normal processing that is manifest in the patient’s clinical signs and symptoms.

This link between symptoms and brain connectivity was well noted by pioneers like Carl Wernicke (1848–1905) who argued that the disorder arose from pathology of the brain’s association fibers. Later, Swiss psychiatrist Eugen Bleuler (1857–1939) who coined the term “schizophrenia” in 1911, defined the illness as a disease whose core involves a de-coupling of the brain’s normally integrated processes [36].

In the last 30 years, the evidence derived from Magnetic Resonance Imaging (MRI), has provided a wealth of evidence indicating that schizophrenia does not solely arise from isolated damage to one or a few brain regions.
Rather, it is characterized by abnormal structural and functional brain connectivity supporting the notion that it is essentially a disorder of dysconnectivity, accounting for a wide-range of connectivity disturbances, beyond a simple “disconnection”, or reduced connectivity [37-41]. A critical point is that these brain connectivity abnormalities are already evident in the initial stages of the disorder and proceed with ongoing illness [42]. This suggests that assessment of brain connectivity patterns with neuroimaging approaches may permit early diagnosis and prediction of disease course.

In recent years, the study of connectivity abnormalities in schizophrenia has been approached not only by examining single functional and structural connections but also from a "connectomic" perspective of neural systems. Connectomics refers to scientific attempts to accurately map the set of neural elements and connections comprising the brain, collectively referred to as the human connectome, at either micro-, meso- or macroscopic resolutions [43].

In this section, we summarize contemporary neuroimaging evidence for abnormal structural and functional brain connectivity, and discuss studies showing aberrant connectome organization in schizophrenia and how these changes have been reported to relate to phenomenology, treatment response and clinical outcome.

Structural Connectivity

Structural connectivity reflects how brain regions are interconnected by white matter fiber bundles and can be assessed using Diffusion-Weighted Imaging (DWI), an imaging method that uses the diffusion of water molecules to generate contrast in MRI [44]. Diffusion Tensor Imaging (DTI), a diffusion-weighted MRI method, can measure macroscopic axonal organization in nervous system tissues by estimating the constraints to diffusion of water molecules in the brain imposed by white matter fiber tracts, allowing these tracts to be delineated [45]. In addition, DTI can be used to approximate the strength or quality of structural connections, by mapping microscopic details about white matter fiber structure [46]. Commonly used metrics of inter-regional connectivity include the number of reconstructed fibers between two regions as a proxy of connection strength [47] as well as measures of connection quality such as Fractional Anisotropy (FA) (how constrained the water diffusion is), Mean Diffusivity (MD) (i.e., the level of water diffusion along each point of a tract), Radial Diffusivity (RD) (i.e., the degree of diffusion perpendicular to the primary tract axis; often taken as an indirect estimate of myelin integrity) and Axial Diffusivity (AD; the degree of water diffusion parallel to the tract trajectory; often taken as an indirect measure of local fiber organization) [48].

DTI Scans can be analysed by different methods: tractography allows calculation and visualization of probable trajectories of fiber tracts with an index of fiber integrity averaged over the reconstructed tract, allowing tract-specific measurements; Voxel-Based Morphometry (VBM) allows DTI metrics to be calculated on a voxel by voxel basis over a normalized stereotaxic brain atlas [46]. Statistical comparison between different populations is then applied to the whole brain or specific regions of interest (ROIs). More recently, Tract-Based Spatial Statistics (TBSS) has been developed in which voxel-based diffusion metrics are projected onto an alignment-invariant tract representation [46].

DTI investigations have reported that changes in white matter may reflect a critical pathophysiological marker for psychosis and schizophrenia. These studies have put in evidence a set of structural brain connectivity impairments in schizophrenia [43, 46, 49]. Altered structural integrity of white matter has been found in major white matter bundles across the brain of schizophrenia patients, especially those connecting the frontal, temporal and parietal lobes, namely the corpus callosum, cingulum bundle, uncinate fasciculus, and internal capsule [49]. These white matter tracts were shown to have reduced fractional anisotropy, believed to be an indicator of affected overall microstructure, together with increased mean and radial diffusivity, which has been suggested to reflect possible decreases in level of myelination and organization of these tracts [49, 50]. In addition to tract-specific differences, diffusion abnormalities have been demonstrated throughout the entire white matter [46], globally suggesting impairment of multiple brain circuits in schizophrenia.

Anticipating the importance of structural connectivity measurements in supporting clinical decision-making in the near future, some studies have reported relationships between structural connectivity alterations and characteristic clinical features of schizophrenia. One of those studies has found that deficits in working memory and reduced speed in information processing, two of the fundamental cognitive deficits in schizophrenia, are associated with impaired whole-brain white matter microstructure [51]. In deficit syndrome schizophrenia, characterized by prominent negative symptoms and poor functional outcome [52], impairments in specific white matter tracts have been found, particularly in the uncinate fasciculus, inferior longitudinal fasciculus, and arcuate fasciculus [53]. Another study, in a more clinically heterogeneous sample of schizophrenia participants, also found that reduced FA in the inferior longitudinal and arcuate fasciculus was associated with functional outcomes at both baseline and 6-month follow-up [54].

A valuable approach that can assist the research for biomarkers in schizophrenia are neuroimaging studies of patients who are experiencing their first-episode of psychosis. Such studies look for white matter abnormalities already present at illness onset. Many of these studies are performed when patients have had none or minimal exposure to anti-

---

[1] The distinction between the prefixes ‘dis’ and ‘dys’ was drawn on etymological grounds by Stephan et al.: in Latin, ‘dis’ refers to a reduction or breakdown; in Greek, ‘dys’ connotes bad or ill and implies that changes could involve both abnormal decreases and/or increases.
psychotic medications. Thus, they can help to distinguish between what is an intrinsic marker of the disorder, and what are secondary effects of long-term medication and chronicity. In first-episode patients evidence exists for tract specific white matter abnormalities in many of the same tracts identified in chronic patients, though findings are less consistent [46]. In contrast, several DTI studies of first episode schizophrenia subjects have reported widespread white matter changes that range from diffusion alterations in frontal, parietal and temporal lobes [55, 56] to an overall reduction in white matter tissue [57]. Conversely other studies have found no differences in all regions assessed in first-episode samples of patients compared to healthy individuals [58-60]. This may be due to some of these studies being underpowered as a result of having small sample sizes as these patients are more difficult to recruit.

Additionally, white matter integrity has been associated to treatment response with first-line antipsychotics in first psychotic episode patients. In a prospective design study, Reis Marques et al. [61] found that non-responders showed lower FA than both responders and healthy control subjects mainly in the uncinate, cingulum and corpus callosum, whereas responders were indistinguishable from healthy control subjects. Interestingly, after 12 weeks of treatment, there was an increase in FA in both responders and non-responders that positively correlated with antipsychotic exposure, raising the possibility that antipsychotics may restore white matter integrity as part of the therapeutic response.

Altogether, studies in first-episode patients suggest that white matter abnormalities are present in the initial stages of schizophrenia and their identification may assist early diagnosis and inform patient stratification to anticipate care needs.

**Functional Connectivity**

Functional connectivity refers to a statistical dependence between spatially distinct neurophysiological signals [62]. This construct assumes that synchronized activity between two or more brain regions reflects some degree of communication occurring between them [50].

Functional Magnetic Resonance Imaging (fMRI) is a form of MRI that enables measurement of brain activity and connectivity [63]. fMRI is commonly performed using the Blood Oxygenation Level-Dependent (BOLD) contrast technique which is a non-invasive imaging technique that indirectly measures brain activation by quantifying magnetic effects of the so-called hemodynamic response. The magnitude of the BOLD signal is an indirect measure of neuronal activity, and is a composite, which reflects changes in regional cerebral blood flow, volume, and oxygenation [64].

fMRI studies typically measure functional connectivity by analysing the distinct neuronal responses to various forms of stimuli and activity during task performance. More recently, technical advances in neuroimaging have revealed intrinsic signal fluctuations during "resting-state", that is, in the absence of external stimuli or demands of imposed tasks [65]. Resting state fMRI (rs-fMRI) studies have been decisive to the recognition that the brain is more accurately portrayed as a network of functionally connected (co-varying) and constantly interacting regions [64, 66].

In schizophrenia, functional imaging studies have found aberrant connectivity within an array of functional brain systems, including fronto-parietal networks involved in cognitive function [67, 68], cingulo-opercular systems associated with salience processing and interoception [69], the default mode network implicated in self-reflective thinking, as well as specific [70] fronto-striatal [71], fronto-temporal [72] and cerebellar [73] circuits. In addition, functional decoupling between distinct subnetworks has been shown in schizophrenia patients [74].

Similarly to structural connectivity, functional connectivity alterations have been reported in association with clinical variables, revealing their potential as candidate biomarkers. In a case-control study of Fornito and colleagues [71] a widespread dysregulation of corticostriatal systems, characterized most prominently by hypo-connectivity of dorsal and hyperconnectivity of ventral frontostriatal circuits, was associated with the early stages of psychosis reflecting differences between first-episode patients and healthy controls. In addition, these changes correlated with the severity of both positive and negative symptoms, and were also present in unaffected first-degree relatives, suggesting that it represents a candidate risk phenotype for psychotic illness. Using resting-state functional MRI, Hadley et al [75] examined the functional connectivity of the ventral tegmental area (VTA) in a sample of unmedicated schizophrenia patients and healthy controls, at baseline and after 1 week of treatment with the antipsychotic drug risperidone. Compared with controls, patients exhibited significantly reduced baseline VTA/midbrain connectivity to multiple cortical and subcortical regions, including the dorsal anterior cingulate cortex (dACC) and thalamus. Importantly, baseline VTA/midbrain connectivity strength to dACC was positively correlated with good response to a 6-week course of risperidone, whereas VTA/midbrain connectivity strength to the default mode network was negatively correlated, suggesting that VTA/midbrain resting-state connectivity may be a useful biomarker for the prediction of treatment response. Moreover, antipsychotic treatment was shown to re-establish deficient baseline VTA/midbrain connectivity to bilateral regions of the thalamus. The effects of treatment in functional connectivity were also examined by Sarpal et al. [76] in a prospective study where the efficacy of treatment of psychosis with second-generation antipsychotic medications was shown to be associated with alterations in functional corticostriatal circuitry. In this study, antipsychotic treatment increased functional connectivity between the striatum and the prefrontal—as well as limbic regions—to the extent that psychosis successfully improved, raising the
hypothesis that increased functional connectivity between these brain regions could be used to predict improvement in symptoms associated with antipsychotic treatment. Building on these findings, in a subsequent study [77], the same authors aimed to develop and test a prognostic biomarker, based on functional connectivity of the striatum, with the potential for clinical utility in the prediction of positive symptom response to antipsychotic treatment. To that end, resting-state functional MRI was used to develop a prognostic baseline striatal connectivity index in a discovery cohort of patients with first-episode schizophrenia, which was then tested in a generalizability cohort of patients with chronic psychotic illness who were newly hospitalized for an acute psychotic episode. This striatal connectivity index predicted response to antipsychotic treatment with high sensitivity and specificity in both the discovery and generalizability cohorts, providing evidence that individual differences in striatal functional connectivity predict response to antipsychotic drug treatment in acutely psychotic patients.

With further development, the evidence from these studies may have the potential to generate diagnostic and prognostic biomarkers with clinical utility and to reduce the overall burden associated with psychotic illnesses.

**Imaging Connectomics and Brain Networks**

The field of 'connectomics' is concerned with accurately mapping the brain's connectivity architecture and understanding how the architectural organization of these neural maps underlies brain function [50]. Imaging connectomics refers to the use of neuroimaging techniques to map various properties of structural and functional brain connectivity, principally at macroscopic resolution [43]. Contributions from network science, mainly based on the mathematical graph theory, have facilitated the study of brain networks and their relation to other complex systems [78-80]. Application of graph theory provides a mathematical framework that describes structural or functional connections between every pair of regions comprising the brain in the form of a graph consisting of a set of nodes (representing neurons, or macroscopic cortical regions) and edges, which describe the interactions between nodes (being axons, macroscopic white matter pathways, level of inter-regional functional coupling) [50] (see Figure 1 for a comprehensive description of the steps involved in the generation of structural and functional brain networks). One the basis of such formal, mathematical description of a network, graph theory metrics can be used to describe the network topology (Figure 2), allowing investigation of emergent features that are otherwise not measurable by focusing exclusively on information from single brain regions or single pair-wise connections.

Based on the type of information the measure provides about the network organization, graph metrics can be approximately divided into measures of segregation and integration. Measures of segregation reflect the degree to which the network can be subdivided into local communities of clusters or modules that are strongly interconnected, with relatively sparse connectivity to the rest of the network. Network clustering can be quantified using the clustering coefficient, indicating the extent to which the neighbours of a node are also mutually connected and therefore have tendency to cluster. Modularity reflects the extent to which the network as a whole can be decomposed into modules.

Measures of integration reflect the efficiency of communication among all nodes in the overall network. Two commonly used and inversely related measures are path length and global efficiency. Path length indicates the average length of the shortest communication paths integrating otherwise spatially disparate node pairs. Global efficiency, calculated as the inverse of the number of steps (i.e. shortest path length) needed to travel between each pair of nodes in the network, reflects the efficiency with which information can be distributed throughout the network (i.e., the degree to which inter-regional communication is globally integrated). Furthermore, degree and centrality metrics provide information on the role of each node and edge in the overall communication architecture of the network [50]. Degree measures the number of edges each node possesses and the level of closeness centrality provides an indication of the travel distance from a node to every other node in the network. In addition, a node's betweenness centrality is a ratio of the number of all shortest paths between any two nodes in the network that travel through an index node, providing an indication of how topologically central a node's role is in overall network communication. Nodes that display a high degree and high of centrality are often referred to as 'hubs' meaning they have a central placement in the network [50, 81, 82] (Figure 2).

Using graph theoretical approaches, studies of the human connectome have identified a non-random organization in the brain's wiring architecture, including a propensity for nodes (i.e., brain regions) to cluster into structural communities [83]. The brain's tendency for local clustering is coupled with a high capacity for global information flow, as indicated by high global efficiency and short path length [78, 83]. This type of network architecture, combining a high level of local information processing while at the same time permitting a high level of global integration of information has been referred to as a small-world topology [83, 84]. The connectome's modular community structure is thought to be advantageous to brain function because high connectivity between nodes in the same module provides a topological foundation for segregation [43, 85, 86] of specialized functions such as primary sensory processing while reducing wiring cost [87]. The structural communities and associated functional systems of the brain have been shown to be interconnected by a relatively small set of putative 'brain hubs' localized in association regions of the frontal, parietal, and insular cortex that display a high pattern of connectivity and a high level of network...
**Figure 1.** Steps for generation of structural and functional brain networks derived from neuroimaging data. (A) Structural connectivity maps can be generated with diffusion tensor imaging (left) and functional connectivity maps can be generated with fMRI (right). (B) Network nodes, corresponding to different brain regions, are identified by parcellation scheme, which can be derived from a priori anatomical landmarks (left), random parcellation (middle) or using functionally defined regions-of-interest (right). (C) After definition of brain regions, structural connectivity between brain regions can be derived from whole-brain tractography for diffusion tensor imaging (left); Functional connectivity is often derived from correlation analysis of statistical dependencies in recorded regional-activity time courses for fMRI (right). (D) After some measure of connectivity has been calculated for every pair of brain regions, connectome architecture can be represented by a connectivity matrix encoding the strength and type of connectivity between each regional pair. (E) The connectivity matrix is used to construct a graph-based representation of brain network connectivity, termed a brain graph (left) from which several measures of network connectivity and topology can be computed (right). fMRI=functional MRI. Panels A-E adapted from Fornito et al. (39)
Biomarkers in Schizophrenia: neuroimaging to machine learning

Figure 2. Graph metrics of network topology. (A) The 'clustering' of a graph provides insight into the level of local connectedness of the graph, describing how strong the connected neighbours of a node are connected themselves. (B) The modular structure of the network describes the tendency of nodes to form local connected clusters that share a relative high level of connectivity with each other than with other regions. (C) The 'characteristic path length' of a graph describes the average number of edges that have to be crossed to travel between any two nodes in the network. (D) The blue node represents a hub node, displaying a high degree, a short global path length and being involved in a large number of communication paths in the network. (E) Besides being individually rich in connectivity, high degree nodes can display a high level of interconnectivity, forming a densely connected 'rich club'.

Numerous topological disturbances of structural and functional brain networks have been found in schizophrenia. Structural studies have generally found evidence of increased segregation (i.e., clustering and modularity) and reduced integration (higher path length and lower global efficiency and rich-club organization), indicating a subtle randomization of network connectivity [50]. In contrast, functional brain network studies paint a more diffuse picture, some reporting reductions in measures of local information-processing such as reduced clustering, local efficiency and/or modularity [88, 89] along with increases in measures of global functional integration others with no significant changes in measures of global integration [88, 90]. Similarly, graph-analysis MRI studies of whole-brain connectivity almost always show impaired structural connectivity [41] while studies of functional connectivity have included reports of both abnormally increased and abnormally decreased functional connectivity, even in the same sample [91]. This apparent discrepancy between findings from structural and functional studies may result from either pathophysiological processes or merely represent an artefact of different methods applied to fMRI and dwMRI data when mapping brain connectivity. Although a detailed discussion on this topic is beyond the scope of this review (for an in-depth overview see [92]), several lines of evidence suggest that a critical factor in determining these topological alterations appears to be a core disturbance of hub node organization schizophrenia [82, 93, 94]. Collectively, these findings indicate that deficient wiring of these hubs and their rich clubs can dysregulate communication across widespread areas, causing complex changes in brain dynamics involving both abnormal increases and decreases in functional connectivity. In some cases, the increases may be a compensatory response to dysregulated signalling in specific parts of the network; in other cases, abnormal wiring of structural connections may lead to a breakdown of normally segregated systems and a de-differentiation of neural activity [92]. Overall affected hub connectivity and topology may underlie a broad range of network abnormal-
ities observed in patients, having a pronounced effect on the functional dynamics and functional capacity of the brain and consequently on global cognitive functioning [50].

Potential of connectomics in informing clinical practice
Beyond demonstrating brain disturbances in schizophrenia, differences in network connectivity or topology may relate to clinical symptoms, treatment response, and functional outcome in schizophrenia.

In patients, positive symptom severity has been linked to reduced levels of overall structural connectivity, both decreases and increases in structural and functional coupling, connectivity strength of temporal and frontal region pairs, reduced global network efficiency [96] and reduced levels of clustering [95, 96]. The severity of negative symptoms has been associated with reductions in global functional connectivity, increases in global structural-functional coupling, increases in structural-functional coupling in default mode subnetworks, as well as reduced levels of global network efficiency and global clustering [95, 96]. Lower functional connectivity has also been associated with poorer cognitive performance [74].

Network-level connectivity analysis may be also valuable for early diagnosis. A recent functional connectivity study that used resting-state functional MRI found that network-level alterations are present in drug-naive individuals with a first-episode of psychosis and suggested that aberrant internetwork connectivity between the default mode network and the central executive network is a distinctive feature of psychosis whereas aberrant intranetwork connectivity is a transdiagnostic feature of psychiatric illness [97].

Concerning treatment response, Crossley and colleagues [98] evaluated connectome-based descriptions of brain structural networks of patients at their first episode psychosis, and found that those who had an efficiently-wired connectome (higher global efficiency) at first onset of psychosis showed a better subsequent response to antipsychotics. However, response to treatment was not accompanied by specific structural changes over time.

In terms of functional performance, impaired rich club organization has been linked to global functioning [99] and is predictive of subsequent changes in functional outcome. Finally, connectome clustering at baseline has been shown to predict longitudinal changes in symptoms and intellectual performance in patients [100].

Taken together these findings suggest that topological alterations in brain wiring have a potential value as markers of disease severity, early detection, treatment response, and outcome prediction.

Machine learning approaches
Neuroimaging studies such as those mentioned until this point have investigated possible biomarkers for early diagnosis, treatment planning and monitoring of disease progression in schizophrenia. To date, however, the results of these studies have had minimal clinical impact despite much interest in the use of brain scans for diagnostic and prognostic purposes [16]. One explanation for the limited impact of the findings on clinical practice is that neuroimaging studies have described average differences between groups whereas decisions in the clinical setting need to be made at the individual level rather than the group. This has been a consequence of relying on mass-univariate analytical techniques (e.g. statistical parametric mapping) which can only be used to detect differences between groups but do not allow statistical inferences at the level of the individual [101]. In addition, with mass-univariate techniques statistical inferences are drawn from multiple independent comparisons (i.e. one for each voxel) based on the assumption that different brain regions act independently, although current evidence argues for a network-level brain function in health and disease [102].

Over the past few years there has been growing interest within the neuroimaging community in the use of analytical methods based in Machine Learning (ML), an area of artificial intelligence that aims to develop algorithms that discover trends and patterns in existing data and use this information to make predictions on new data.

Relative to traditional methods of analysis based on the mass-univariate analytical techniques, the advantages of applying ML methods are twofold. Firstly, ML methods allow statistical inferences at the level of the individual therefore yielding results with a potentially high level of clinical translation [16]. Secondly, as inherently multivariate approaches, they take the intercorrelation between voxels into account, and thereby are sensitive to spatially distributed and subtle effects in the brain that would be otherwise undetectable using traditional univariate methods [101].

ML methods can be divided into two broad categories: supervised and unsupervised learning. In supervised ML, one seeks to develop a function which maps two or more sets of observations to predefined categories or values. In contrast, unsupervised methods seek to determine how the data are organized without using any a priori information supplied by the operator; here the main objective is to discover unknown structure in the data [101].

One specific form of supervised multivariate pattern recognition analysis (MVPA) algorithm is that used for classification, concerned with the automatic discovery of regularities in the data that can be used to classify the data into different predefined categories [16]. Examples of MVPA techniques include, but are not limited to, artificial neural networks, decision trees, Gaussian process classification and Support Vector Machine (SVM). Of these, the most popular technique in the neuroimaging literature is SVM [103], which to date has been applied in several studies of neurological and psychiatric disorders, allowing the classification of individual
observations (e.g. scans) into distinct groups or classes (e.g. diagnostic categories) based on data in high-dimensional space [104].

The SVM is a supervised classification algorithm that learns from an initial ‘training’ data set to classify new cases into two or more groups previously defined by the operator (e.g. patients vs. controls) [16]. For example, an SVM can be trained on a data set of baseline neuroimaging measurements from first-episode subjects defined by whether they subsequently developed a clinically defined outcome or not, and it can then be applied to a new sample of first-episode individuals [16]. An SVM is validated by demonstrating that it can classify individuals in a sample independent of the data set it was originally trained on. Ideally, this is done using a second data set that has been acquired separately from the original [16]. Validation can also be attempted without a second data set, for example, by splitting the original sample (repeatedly), with one portion being used as the training data set and other portions being used for cross validation (e.g., k-fold cross-validation; [105]). However, any overlap between the training and test data sets will result in the model over-fitting.

Several studies have used ML to examine the diagnostic and prognostic potential of neuroimaging in a range of neurological and psychiatric disorders and, to date, a number of promising results have been reported. In schizophrenia, a meta-analysis of studies that applied MVPA to brain functional and structural imaging data revealed that this ML method differentiates schizophrenic patients from healthy controls with 80% sensitivity and specificity [14]. Another potential application of MVPA is in the field of differential diagnostics. In fact, because of the substantial symptomatic heterogeneity within and overlap across different psychiatric diagnoses, differential diagnostics might represent the greatest clinical challenge in everyday care [106]. Thus, MRI based differential diagnosis may develop into one of the most promising applications of MVPA. In recent study, structural MRI-based multivariate pattern classification was used to identify and cross-validate a neuroanatomical signature that allowed differential diagnosis of patients with first-episode and recurrent stages of schizophrenia from patients with major depression with a balanced accuracy of 76% [107].

Results from the application of SVMs in first-episode samples have also been promising. A SVM trained on volumetric MRI data obtained at the first psychotic episode was used to predict future illness course at the individual level [108]. The trained SVM was able to distinguish patients with a continuous course from both patients presenting an episodic course (accuracy=70%, sensitivity=71, specificity=68, p=0.004) and healthy individuals (accuracy=67%, sensitivity=71, specificity=61, p=0.01). This study provided novel, methodologically sound, proof of concept that ML methods applied to MRI data can be used at illness onset to predict clinical outcome.

Overall, findings show that application of ML methods to neuroimaging data may provide generalizable tools for diagnosis and improved treatment-management of schizophrenia early in the course of the disorder.

Concluding remarks and directions for future research

In recent years, precision medicine has become a ubiquitous paradigm across all areas of medicine, including schizophrenia, boosting research for personalised care. The neuroimaging studies reviewed above support the perspective that use of imaging methods that assess neurochemical brain functioning such as dopamine PET or [1H]-MRS or newer methodologies measuring functional or structural connectivity, may be suitable to discover biomarkers of the disorder.

An important step in the research for biomarkers in schizophrenia is the identification of pathological phenotypes that are characteristic of the disorder. However, many of the pathological features of schizophrenia may be elusive to "resting-state" neuroimaging methods, requiring challenge paradigms to be revealed. As such, task-based studies should be considered crucial for detecting the characteristic features upon which biomarkers can be developed. In addition, the identification of new candidate biomarkers may require the use of several imaging modalities. Indeed, schizophrenia is a heterogeneous and complex disorder and biomarkers of a single imaging modality do not capture the heterogeneity and complexity of this brain disorder. Until recently, integration of different imaging modalities into a single candidate biomarker represented a significant methodological challenge, since multimodal studies require broader expertise in collecting, analysing, and interpreting the results than do unimodal studies [109]. Nevertheless, the recent development of toolboxes designed to pipeline well-known neuroimaging software, allows automated connectivity analyses and extraction of graph theory metrics from multimodal imaging data such as PET, MRS, fMRI and DTI [110]. In addition, the application of new machine-learning methods using multivariate approaches to fuse brain imaging data from different modalities and non-imaging data has the potential to identify endophenotypes and increase knowledge on individual diagnosis and treatment response [109, 111, 112].

Since the final purpose of developing neuroimaging biomarkers is to provide tools to aid real-world clinical management of schizophrenia patient, neuroimaging findings can be developed as biomarkers either with diagnostic purposes by indexing a biological process associated with health or the disorder or as predictive biomarkers reflecting a process associated with treatment-response, that can be used in clinical stratification [113]. However, a neural abnormality found in a clinical population will only be
assigned as a real biomarker if it can be used as an accurate surrogate of clinically relevant outcomes, such as diagnosis, prognosis or treatment response, which means it needs to have sufficient sensitivity, specificity and clinical predictive value. One of the first obstacles for determining the predictive power of biomarkers for schizophrenia relates to the lack of a gold standard definition for diagnosis. The current diagnostic criteria are solely based on combinations of symptoms and there are no gold-standard objective methods that can be used for definitive validation of clinical diagnosis. A possible solution has been to conduct longitudinal follow-up studies, which are considered equivalent to a gold standard in psychiatry and allow comparison of the biomarker prediction with the final diagnosis or observed clinical outcome. An alternative approach could be the use of new classification schemes such as the Research Domain Criteria (RDoC) initiative, which is based on dimensions of observable behavior and neurobiological measures, rather than categorical diagnosis as in the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Classification of Diseases (ICD). The RDoC approach can aid the search for biomarkers because it might lead to the identification of neuroimaging markers of discrete symptom domains of schizophrenia, which then can be targeted by specific treatments instead of treating the disorder as a whole.

A major limitation for the development of biomarkers in schizophrenia has been the relatively small sample sizes of studies published until now, resulting in lack of statistical power to allow validity assessment of suggested candidates. Systematic imaging of a critical number of patients may unveil new biomarkers that otherwise would go undetected in underpowered studies. A promising strategy to enable well-powered studies could emerge from precise-medicine initiatives like the PSYSCAN project (http://www.psycan.eu/), a multi-centre study funded by the European Commission, involving a large number of patients, aimed at developing a neuroimaging-based tools that can be used to help predict outcomes in the early phase of schizophrenia. Of course, multi-site studies are challenging, as they require standardization of neuroimaging methods in order to ensure reproducibility across sites. This could be tackled with standardized protocols that provide homogeneous data acquisition and analysis across sites, including procedures for subject preparation, performance of quality-assurance tests on the scanner before data acquisition is carried out, as well as specific analysis software for data processing [113].

Finally, a practical issue that can limit the clinical use of imaging biomarkers is the restricted availability of imaging centres and the substantial costs associated with brain imaging. Nevertheless, the cost-effectiveness defined by the clinical usefulness of the biomarker in avoiding expenses related with misdiagnosis, treatment failure, functional impairment and other disease-related burden may be more important than cost [113].

In sum, for a biomarker to be translated as a useful tool for clinical management of schizophrenia patients, it is required to have an acceptable level of sensitivity, specificity and predictive value, to be easily quantifiable and practically feasible, easily accessible and cost effective.

In the future, biomarker development in schizophrenia should rely on a cost-effective platform that balances between hypothesis-driven development, exploring well-replicated imaging phenotypes with the potential to result in useful biomarkers, and data-driven approaches aimed at ‘blindly’ exploring multimodal data sets [113]. Ultimately, once developed, these methods could allow a patient to be assigned to targeted assertive case-management at first presentation to services, with optimized pharmacological treatment, cognitive and family interventions, which have been shown to improve treatment adherence and reduce relapse rates eventually improving outcome [108]. This means that the discovery of neuroimaging-based biomarkers of schizophrenia would represent a paradigm shift in patient care, positively influencing the social perception of mental illness and the quality of psychiatric practice.

**Abbreviations**

[1H]-MRS: Proton Magnetic Resonance Spectroscopy; [18F]-DOPA - [18F]-L-6-fluoro-3,4-dihydroxyphenylalanine; AD: Axial Diffusivity; BOLD: Blood Oxygenation Level-Dependent; D2: Dopamine Type-2 Receptor; D3: Dopamine Type-3 Receptor; dACC: Dorsal anterior cingulate cortex; DSM: Diagnostic and Statistical Manual of Mental Disorders; DTI: Diffusion Tensor Imaging; DWI: Diffusion-Weighted Imaging; FA: Fractional Anisotropy; fMRI: Functional Magnetic Resonance Imaging; GABA: Gamma-Aminobutyric Acid; Gln: glutamine; ICD: International Classification of Diseases; MD: Mean Diffusivity; ML: Machine Learning; MRI: Magnetic Resonance Imaging; MVPA: Multivariate pattern recognition analysis; NMDA: N-Methyl-D-Aspartate; PET: Positron Emission Tomography; rs-fMRI: Resting state fMRI; RD: Radial Diffusivity; RDoC: Research Domain Criteria; ROIs: Regions of interest; SPECT: Single Photon Emission Computed Tomography; SVM: Support Vector Machine; TBSS: Tract-Based Spatial Statistics; VBM: Voxel-Based Morphometry; VTA: Ventral tegmental area.

**Competing interests**

The authors declare no conflict of interest.

**References**


Biomarkers in Schizophrenia: neuroimaging to machine learning

https://doi.org/10.1002/j.2051-5545.2011.tb00101.x


37. Friston KJ. The disconnection hypothesis. Schizophrenia research 1998; 30(2):115-25. https://doi.org/10.1016/S0920-9964(97)00140-0


65. Fox MD, Raichle ME. Spontaneous fluctuations in brain activity
observed with functional magnetic resonance imaging. Nat Rev Neurosci 2007; 8(9):700-11. https://doi.org/10.1038/nrn2201


109. Calhoun VD, Sui J. Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness. Biological psychiatry Cognitive neuroscience and neuroimaging 2016; 1(3):230-44. https://doi.org/10.1016/j.bpsc.2015.12.005


