Do we have any solid evidence of clinical utility about the pathophysiology of schizophrenia?

Citation for published version:
Lawrie, SM, Olabi, B, Hall, J & McIntosh, AM 2011, 'Do we have any solid evidence of clinical utility about the pathophysiology of schizophrenia?' World psychiatry : official journal of the World Psychiatric Association (WPA), vol. 10, no. 1, pp. 19-31.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
World psychiatry : official journal of the World Psychiatric Association (WPA)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Do we have any solid evidence of clinical utility about the pathophysiology of schizophrenia?

Stephen M. Lawrie, Bayanne Olabi, Jeremy Hall, and Andrew M. McIntosh

Division of Psychiatry, School of Molecular and Clinical Medicine, Royal Edinburgh Hospital, Morningside, Edinburgh EH10 5HE, UK

A diagnosis of schizophrenia, as in most of psychiatric practice, is made largely by eliciting symptoms with reference to subjective, albeit operationalized, criteria. This diagnosis then provides some rationale for management. Objective diagnostic and therapeutic tests are much more desirable, provided they are reliably measured and interpreted. Definite advances have been made in our understanding of schizophrenia in recent decades, but there has been little consideration of how this information could be used in clinical practice. We review here the potential utility of the strongest and best replicated risk factors for and manifestations of schizophrenia within clinical, epidemiological, cognitive, blood biomarker and neuroimaging domains. We place particular emphasis on the sensitivity, specificity and predictive power of pathophysiological indices for making a diagnosis, establishing an early diagnosis or predicting treatment response in schizophrenia. We conclude that a number of measures currently have the potential to increase the rigour of clinical assessments in schizophrenia. We propose that the time has come to more fully evaluate these and other well replicated abnormalities as objective potential diagnostic and prognostic guides, and to steer future clinical, therapeutic and nosological research in this direction.

Key words: Schizophrenia, etiology, pathophysiology, diagnosis, early diagnosis, treatment response, predictive power, likelihood ratio

(World Psychiatry 2011;10:19-31)

METHODS OF THE REVIEW

As we are interested here in clinical utility over and above statistical significance, we concentrate on studies which provide data in terms of the sensitivity and specificity of the variables as a diagnostic aid, the predictive power of a test result and/or the likelihood that a test result in an individual patient is indicative of schizophrenia. It is worth noting that sensitivity and specificity are generally constant properties of a test, which are useful in service planning but not in dealing with individual patients. The positive predictive value (PPV) or negative predictive value (NPV) of a test result gives the risk level for a particular patient, which is useful clinically, but PPVs and NPVs are prevalence-dependent measures, and performance can therefore vary markedly in different settings (11). Likelihood ratios are a means of using sensitivity and specificity data to calculate the implications of test results in a particular patient (12-14). As a rough rule of thumb, likelihood ratios of a positive test result (LR+) of more than 5, and preferably more than 10, increase the risk of disorder by about 50% or 45%, respectively. The latter would, for example, indicate a clear change from a pre-test probability of say 50% (maximal
uncertainty) to a post-test probability of 95% (highly likely). This might at first appear to be an alien practice, but it is for example what underpins the use of the CAGE questionnaire in identifying alcohol problems and the Mini Mental State Examination in diagnosing dementia (15,16).

The type of study we need for a diagnostic test is a cross-sectional one comparing a representative population of patients and non-cases (controls for diagnosis, other diagnoses for differential diagnosis) who have been evaluated with the gold standard and blindly assessed with the test. For early diagnosis and treatment response tests, we need a longitudinal and preferably prospective study of a cohort of patients evaluated before or after the onset of their condition and followed up until outcome is clear, with preferably less than 20% loss to follow-up.

In this review, we sought to identify replicated evidence from systematic reviews for the diagnosis, early diagnosis and treatment response of schizophrenia, in terms of the reliability of the examination, the size of the difference between schizophrenia and controls, and the ability to discriminate versus bipolar disorder. In each of our specified domains, we particularly sought reviews with some consideration of measurement reliability, heterogeneity and publication bias. We favoured reviews reporting an effect size such as Cohen’s d of 1 or more, as this roughly and generally corresponds to a 70% non-overlap of data distributions and an odds ratio (OR) of approximately 5 (17).

DIAGNOSIS

In everyday medical practice, history-taking identifies diagnostic “hypotheses”. Evidence for and against these is sought on physical examination and (ideally) confirmatory diagnostic testing. In psychiatry, a similar initial approach is followed by the mental state examination, which includes more explicit evaluation of appearance, behaviour and speech than in the rest of medicine, but also several questions that are just further history gathering and some cognitive testing often of dubious validity. We psychiatrists are curiously averse to physically examining our patients and surprisingly willing to accept “CNS grossly normal” in medical records when this probably means that no neurological exam has been attempted. We might consider possible “organic” explanations for “secondary schizophrenia” and contemplate referral for brain imaging in unusual cases, but that is about as far as investigation is usually taken.

To provide some clinical background, and as a comparator for laboratory tests, we first consider the evidence base for key aspects of the clinical examination in making a diagnosis of schizophrenia.

Clinical history and examination

Psychotic symptoms

Although counter-intuitive, and despite the potential tautology, particular psychotic symptom types are not in themselves strong associates of schizophrenia. Bizarre delusions, for example, are less reliably elicited (mean kappa across studies 0.5 or “moderate”) than delusions in general (0.7 or “substantial agreement”) (18), and have a low specificity, despite a PPV as high as 0.82 in 214 consecutive admissions (19). Similarly, Schneider’s FRS have been over-sold as pathognomonic, as they are both too rare to be a generally useful diagnostic aid, especially if strictly defined, and too common in other psychotic disorders (20). Peralta and Cuesta (21) recruited 660 inpatients with “the full spectrum” of psychotic disorders and found that any individual FRS usually had a LR+ of 1-2 for schizophrenia, depending on the particular symptom and the diagnostic criteria examined, and no FRS had a LR+ more than 4.

Risk factors as diagnostic aids

There is clear evidence that several variables increase the risk for schizophrenia to a statistically significant degree (7,8,22). Table 1 lists some of these. Indeed, many of these risk factors, and especially family and developmental history, are sometimes used as supportive evidence to make a diagnosis of schizophrenia, but in an informal and variable way. These factors rarely elevate the risk by more than 5x relative to the baseline population risk of approximately 1%. Even elevating that risk to approximately 10% in the presence of a positive family history in a first-degree relative (10) is clearly not very helpful, and reliably eliciting the information might require structured assessments (23).

Where the presence of such risk factors might be helpful is in differential diagnosis, perhaps particularly in hospital settings where psychosis is much more prevalent, as the major psychoses may breed partially true (24), and urban birth and developmental disruption may be more potent risks for schizophrenia than bipolar disorder (22).

On the other hand, although the risk of schizophrenia is clearly elevated by experiencing obstetric complications (OCs), the additional risk from any one complication is much smaller, and OCs probably increase the odds of a range of adverse neurodevelopmental outcomes (25,26). If one was to make clinical use of the association between immigration and schizophrenia, despite the heterogeneity (27), one might quickly run into allegations of racism. Regular cannabis use is a risk factor for schizophrenia (28,29), but it is sometimes argued,

Table 1 Best replicated historical risk factors for schizophrenia (adapted from 7,22)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level of risk</th>
<th>Key supporting reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family history</td>
<td>RR up to 50</td>
<td>Gottesman (10)</td>
</tr>
<tr>
<td>Immigrant status</td>
<td>OR = 5</td>
<td>Cantor-Graae and Selten (27)</td>
</tr>
<tr>
<td>Childhood social difficulties</td>
<td>OR up to 5</td>
<td>Tarbox and Pogue-Geile (110)</td>
</tr>
<tr>
<td>Obstetric complications</td>
<td>OR = 2-3</td>
<td>Cannon et al (26)</td>
</tr>
<tr>
<td>Cannabis use</td>
<td>OR = 2-3</td>
<td>Moore et al (29)</td>
</tr>
<tr>
<td>RR - risk ratio; OR - odds ratio</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

World Psychiatry 10:1 - February 2011
without much evidence, that this may not be a causal relationship, i.e. that people with pre-schizophrenia take up cannabis use perhaps in a bid to self-medicate (30). Nevertheless, it is clear from randomized controlled trials that cannabinoids prescribed, for example, as anti-emetics for people with cancer increase the risk of hallucinations about six-fold and delusions more than eight-fold (31). Cannabis may therefore only induce psychotic symptoms, and some additional factor or at least chronic, frequent use may be required before schizophrenia develops. Thus, the standard clinical practice of making a diagnosis of drug or cannabis induced psychosis in regular drug users, and watchful waiting to see if schizophrenia develops, is probably rational. We are not aware, however, of any studies examining this practice, or indeed the relative merits of subjective versus objective assessments of cannabis use in so doing.

It should hopefully be clear that we do use risk factors in making a diagnosis of schizophrenia, but as currently implemented this is haphazard.

Physical signs

Despite our reluctance to examine our patients physically, it is clear that there are some physical signs which are risk factors for schizophrenia and of potential pathophysiological significance. Minor physical “anomalies” like abnormal head circumference, hypertelorism, and non-right handedness are, however, too non-specific and only raise the risk of schizophrenia slightly (32,33), while dermatoglyphic patterns are difficult to discern (34). Neurological “soft” signs (NSS) are more promising, as it has been reported that 50-60% of patients with schizophrenia have observable deficits in sensory integration and motor coordination, as compared to about 5% of normal controls (35). In a thorough recent systematic review and meta-analysis, Chan et al (36) found an overall effect size of 1.08, corresponding to a 73% separation (17) between the populations, although this effect size is probably inflated by the difficulties in blinding such assessments to patient and control status. There was, however, largely unexplained statistical heterogeneity, and evidence of publication bias, potentially attributable to difficulties in reliably eliciting these appropriately named phenomena.

Rigorous evaluation of NSS may be particularly difficult in patients with the most acute psychoses. Further, while it is clear that these signs are not simply attributable to antipsychotic treatment, it is unclear to what extent they reflect the nature of the underlying pathophysiological processes of schizophrenia, as disease specificity has only been studied rarely (35). Given, however, that certain NSS may be more genetically mediated (37), and that NSS have been proposed as clinical and functional outcome predictors (35), the area does look promising for further clinically oriented research (see Table 2). There may be value in considering the reliability and diagnostic specificity of individual signs within the major domains of NSS and their likely anatomical underpinnings – motor dexterity (cerebellum), primitive reflexes (frontal lobe), motor sequencing (prefrontal cortex, PFC), and sensory integration (parietal lobe) – in more detail than just a global NSS score.

Cognitive testing

Examining cognitive status in everyday psychiatric practice is usually done with a few quick tests of largely unproven reliability and validity. Evaluating cognitive performance rigorously is not routine outside a research setting, but patients with schizophrenia certainly have a range of intellectual impairments (38), most of which are evident at first episode (39). Meta-analyses have identified large (d>1) deficits in general intelligence (38-40), processing speed (41), various aspects of memory (38,39,42,43), verbal fluency (44), social cognition (45) and theory of mind (46). It remains difficult, however, to establish whether there are specific deficits over and above general performance decrements. There is also marked heterogeneity between studies, potentially attributable to the effects of mental state on performance and cooperation, the fact that many patients can approach normal performances at times, as well as variation in the populations studied and in how assessments are conducted and scored.

From a pathophysiological point of view, there is also the problem that many of these cognitive deficits appear to be largely present prior to the onset of psychosis, with some further deterioration, in some cases at least, after onset; all probably confounded by other risk

Table 2 Large consistent effects from meta-analyses of studies of physical and cognitive examinations of patients with schizophrenia versus controls

<table>
<thead>
<tr>
<th>Effect size vs. controls</th>
<th>Different from relatives</th>
<th>Evident at first episode</th>
<th>Specificity vs. bipolar disorder</th>
<th>Other issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSS</td>
<td>1.08, but with heterogeneity (Chan et al, 36)</td>
<td>Yes</td>
<td>Yes</td>
<td>Requires more study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Blinding reliability and practicality issues; specific domains and items may have stronger diagnostic properties</td>
</tr>
<tr>
<td>IQ</td>
<td>1.10, but with heterogeneity (Heinrichs and Zakzanis, 38)</td>
<td>Yes</td>
<td>Yes, at least in part, but some possible progression</td>
<td>Premorbid IQ deficits may differentiate</td>
</tr>
</tbody>
</table>

NSS – neurological “soft” signs

21
From a clinical perspective, most task deficits also appear to be evident in patients with bipolar disorder and psychotic depression, albeit to a slightly lesser extent (50-52). General intellectual impairments are, however, more commonly found in schizophrenia than bipolar disorder, especially before diagnosis (38,51). It may be that IQ level, and perhaps especially pre- to post-morbid deterioration, would provide useful information in making diagnoses (53,54), or perhaps in identifying a subgroup at risk of poor prognosis and/or in need of aggressive treatment. Given the heterogeneity in IQ assessments in schizophrenia, it may also make sense to evaluate discrete aspects such as processing speed or verbal fluency, perhaps as part of brief assessments with well-evaluated psychometric properties, such as the Brief Assessment of Cognition in Schizophrenia (55).

**Blood tests for “biomarkers”**

**Genomics**

It is well known that schizophrenia has a large heritable component. Genetic factors and gene-environment interactions contribute up to 80% of the liability to the illness (10,56). As the clinical phenotype is complex, and the pathophysiology is likely to be polygenic, the genes involved have been hard to find. In recent years, a number of convergent findings in linkage (57), association and animal studies have consistently implicated several genes, for which the most consistent evidence is arguably for the “Icelandic haplotype” in the neuregulin-1 gene (58), although the overall OR of about 2 and continuing uncertainty about which particular genotype is implicated mean that this remains of purely research interest. The recent complete mapping of the human genome has enabled several genome wide association studies in schizophrenia, which have been meta-analysed to reveal multiple small effects across the genome, with the strongest overall effect (OR ~ 1.09) being in the ZNF804A gene encoding a putative zinc finger protein (59).

Rare variants conferring risk for schizophrenia have also been identified. Perhaps the most striking example is the DSCR1 (Disrupted in Schizophrenia 1) gene, identified in a large Scottish family in which a chromosomal translocation is associated with a high frequency of schizophrenia (60), although this translocation is possibly unique to this family and raises the risk for bipolar disorder and depression as well. Smaller chromosomal abnormalities, known as copy number variants (CNVs), are also more common in patients with schizophrenia than controls. One relatively common example is the 22q11 deletion known to occur in velo-cardio-facial syndrome, which is associated with a greatly increased risk (RR ~ 30x). Notably, this genomic region includes the catechol-O-methyltransferase (COMT) gene, involved in dopamine metabolism, which may also be a risk gene for schizophrenia, perhaps especially in multiply affected families. Initial genome-wide studies of CNVs provide replicated associations of schizophrenia with rare 1q21.1, 15q11.2 and 15q13.3 deletions. Collectively, several rare CNVs may elevate risk for schizophrenia, perhaps especially the more developmental forms of the disorder, but large CNVs do not appear to be implicated in bipolar disorder. Including 22q11.2 deletions, CNVs appear to account for up to 2% of schizophrenia (61). It would however be premature to routinely screen patients for CNVs, both because causality has yet to be established and the information gained might not influence management.

**Proteomics**

Quantitative and qualitative protein patterns in cerebrospinal fluid (CSF) and serum have potential as diagnostic and prognostic biomarkers in schizophrenia and other psychiatric disorders (62-64). There has been much interest in serum brain-derived neurotrophic factor (BDNF) levels in patients with schizophrenia, as BDNF has roles in neuronal proliferation, differentiation and dopamine neurotransmission, but extremely mixed results have been reported compared with controls. Inconclusive results have also been reported for serum levels of epidermal growth factor. There are more consistent results from several studies supporting an association between schizophrenia and S100B, a calcium-binding protein produced primarily by astrocytes, where increased concentrations likely result from astrocyte destruction. Most studies report increases in serum and CSF S100B concentrations in schizophrenia (65-68).

The potential importance of immunity in the pathogenesis of schizophrenia is supported by findings of altered serum concentration of several proinflammatory cytokines. Potvin et al (69) examined data from 62 studies, involving a total of 2298 schizophrenia patients and 1858 healthy volunteers, and found consistent increases in interleukin 6 (IL-6), soluble IL-2 receptor, and IL-1 receptor antagonist, and a decreased in vitro IL-2 in schizophrenia. IL-6 is, however, also reduced in depression, and stress and weight gain are potential confounders (70). This highlights the care required in interpreting these studies, particularly given the infamous “pink spot” in the urine of patients with schizophrenia in the 1960s and the consistently reduced levels of platelet monoamine oxidase (MAO) in the 1980s, which were eventually related to smoking status (71).

**Brain imaging investigations**

There is overwhelming evidence for a variety of consistent abnormalities of brain structure and function and electrophysiology in patients with schizophrenia compared to healthy controls (72,73) (see Table 3 for examples). There are similar concerns as with the cognition studies about when these abnormalities develop. The imaging literature
Table 3 Large consistent effect sizes from meta-analyses of brain imaging studies in patients versus controls

<table>
<thead>
<tr>
<th>sMRI regional brain volumes</th>
<th>Effect size vs. controls</th>
<th>Different from relatives</th>
<th>Evident at first episode</th>
<th>Specificity vs. affective disorder</th>
<th>Other issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypofrontality</td>
<td>0.64 at rest; 1.13 when active</td>
<td>Yes</td>
<td>Yes</td>
<td>Amygdala volume may discriminate but may depend on age and treatment</td>
<td>Pattern recognition methods may be more powerful</td>
</tr>
<tr>
<td>Mismatch negativity</td>
<td>0.99</td>
<td>Possible</td>
<td>Possible, but some possible progression</td>
<td>Possible</td>
<td>-</td>
</tr>
</tbody>
</table>

sMRI – structural magnetic resonance imaging; DLPFC – dorsolateral prefrontal cortex

shows, however, less evidence of heterogeneity across studies and somewhat greater evidence for specificity versus bipolar disorder.

**Structural brain imaging**

Structural magnetic resonance imaging (sMRI) is relatively straightforward, cheap and available, and shows perhaps the greatest current promise as an objective diagnostic test for schizophrenia. The effect sizes are small, but the measures are inherently quantitative. Perhaps the greatest single demonstration of the power of this approach is from the landmark finding that monozygotic twins with and without schizophrenia could be discriminated by simply eyeballing their sMRI scans, and in particular the ventricles and medial temporal lobes, in 80% or more of the 15 pairs (74). Of course, twins are in short supply in clinical practice. More realistic is to use the evidence from what is now a large sMRI literature in schizophrenia, that there are consistent if relatively small reductions in whole brain, PFC and temporal lobe volumes (d = 0.2-0.4), and consistently reduced amygdala volumes (d = 0.7) in schizophrenia (75-77). Moreover, the sMRI changes in schizophrenia are less marked in relatives and others at high risk, show evidence of changes around the time of onset and are largely evident at first episode (78). The effect sizes are greater in schizophrenia than bipolar disorder (79,80), and the amygdala may actually be large or normal in bipolar disorder (79), perhaps particularly in younger patients. This merits intensive study as a possible discriminator, although there are technical difficulties in reliably extracting volumes in such a small structure.

A number of automated support vector machine (SVM) analyses have recently been applied to sMRI data in schizophrenia (81). Generally, 80-90% of patients can be identified from their similarity to a group pattern for schizophrenia (82-84), although these studies do tend to rather circularly use group differences to inform the group classification, and do not convincingly agree on the anatomical patterns of differentiation. The challenges for such studies are to distinguish schizophrenia from bipolar disorder, to generate individual scan readings, to cross test various models on various software routines, and to compare them with other diagnostic techniques including other approaches to brain imaging.

**Functional brain imaging and electrophysiology**

**Hypofrontality**

The relative underactivation of PFC is one of the most consistent findings in schizophrenia research. Zakzanis and Heinrichs (85) found an overall effect size from 21 resting positron emission tomography (PET) studies of -0.64, a 60% overlap in data distributions, and an even greater effect from 9 activated PET studies of -1.13, a 40% overlap, although they did not examine for heterogeneity or publication bias. As currently reported, functional MRI (fMRI) studies do not lend themselves to the calculation of overall effect sizes, but hypofrontality is clearly evident in dorsolateral PFC on working memory studies (86) as it is in (left) inferior PFC on verbal memory tasks (87). Functional imaging studies and especially fMRI also tend to be analysed in relative rather than in absolute terms, preferable for diagnostic evaluations. Nonetheless, several classification studies have found that dorsolateral PFC activation on various tasks might distinguish schizophrenia from bipolar disorder (88,89), and similarly high diagnostic accuracy (>80%) has been reported in default-mode network activity (90) and on resting fMRI (91). However, a recent study found less discrimination, perhaps because task performance differences clouded the picture (92). It remains to be seen how such an approach would cope with the most difficult differential, i.e. when those with bipolar disorder in the sample are experiencing active psychotic symptoms.

**Positron emission tomography (PET)**

PET has also been used to assay neurotransmitter receptors in vivo, and dopamine D2 receptors in particular. This has been a controversial field, but D2 receptors are increased overall, with an effect size of 1.47 across 17 post-mortem and PET studies (93), including some medication naive subjects. Furthermore, there is a very consistent literature dem-
onstrating increased pre-synaptic activity in the striatum, as indexed by greater amphetamine-promoted dopamine release and greater F-DOPA uptake in schizophrenia (94). A preliminary classification study is also encouraging (95), although making a distinction between schizophrenia and bipolar disorder with psychotic symptoms is arguably unlikely.

Electrophysiology

A small number of studies have provided data on the sensitivity and specificity of EEG findings in the differential diagnosis of schizophrenia, with very mixed results (96). Several measures of neuronal responses to stimuli, especially the P300 and P50, show large effect sizes versus controls, but large amounts of unexplained heterogeneity between studies (97). They also tend to show almost as large effects in relatives, suggesting a greater loading on trait rather than state effects, and possibly less utility for diagnosis. Mismatch negativity does, however, show promise in these regards (see Table 3) and has possible specificity (98). Finally, a solitary but impressive study has considered exploratory eye movements in 145 patients with schizophrenia from seven World Health Organization collaborative centres and found more than 85% sensitivity and specificity against depression and healthy controls (99), although a recent Japanese multisite study was less successful (100).

EARLY DIAGNOSIS

Diagnoses have value for communication and prognostication, but particularly for planning action. Early diagnosis is actually akin to accurate prognostication within a group as to who will develop the disorder of interest and who will not. Studies of early diagnosis therefore require lengthy follow-up, and any predictors should ideally be unambiguously defined and measured, and improve upon what can be achieved in current practice. Again, therefore, we first consider the potential role of psychotic symptoms in early diagnosis.

Clinical features

Psychotic symptoms as predictors

A range of childhood psychopathologies have been shown to predict schizophrenia. The strongest of these have included: self-reported psychotic symptoms at age 11, which increased the risk 16x of schizophrenia disorders at age 26 (101); schizophrenia spectrum personality disorder (PD) at mid-teens in Israeli army conscripts males, increasing the risk of schizophrenia by 21.5 times (102); and diagnoses of alcohol abuse, any PD, or substance abuse in Swedish army conscripts aged 18 or 19, increasing the risk of subsequent schizophrenia (OR 5.5, 8 and 14, respectively) (103). These statistical effects are, however, insufficiently replicated and too prone to high false positive rates for clinical use.

Early diagnosis becomes more practically and ethically straightforward when people present as patients with prodromal symptoms. Klosterkotter et al (104) followed 160 prodromal patients over a decade and found that ten “basic symptoms”, including subtle disturbances of mental life such as stress sensitivity, had PPVs of more than 70%. This has yet to be replicated, however. The most common approach has been to use the ultra high risk (UHR) criteria devised as a means to predict transition to psychosis in clinic attenders in Melbourne (105). The transition rates to psychosis (not just schizophrenia) were as high as 54% within 12 months at first, with PPV/NPV both more than 80% (106), but these figures have steadily fallen with time and application in different settings, so that transition rates can now be as low as 14% after 12 months and 19% after 18 months (107).

Several prospective cohort studies have followed children or adolescents at high genetic risk as they are offspring or otherwise related to patients with schizophrenia. Thought disorder and negative symptoms, behavioural or neuromotor dysfunction, and attention and memory impairment are fairly consistent predictors in these studies (108), but only two studies have reported data in terms of clinical prediction. In the New York High Risk Project (NYHRP), the predictive power of symptoms for adulthood schizophreniform psychosis was not that high (107). In the Edinburgh High Risk Study (EHRIS), in which the baseline risk of transition to schizophrenia was 21/162 (13%), psychotic symptoms at interview only had a PPV of 25%, a schizotypal PD at interview only had a PPV of 29% and the strongest behavioural predictor of any sort was a self-completed questionnaire for schizotypal traits (the Rust Inventory of Schizotypal Cognitions, RISC, PPV 50%). All of the foregoing did, however, have NPVs more than 90%, and the RISC figures correspond to an LR+ of >5 (109).

Risk factor prediction

These prospective cohort studies of young people at high genetic risk have also established a number of behavioural abnormalities in childhood and adolescence that predict subsequent psychosis, usually with greater power than family history, migration, OCs or regular cannabis use (108). In the EHRIS, none of those risk factors was a statistically significant predictor of schizophrenia, but several aspects of childhood behaviour, as elicited from the mothers with the Achenbach scale, were (109). Tarbox and Pogue-Geile (110) recently summarized this literature and concluded that “poor underdifferentiated social functioning” is a moderately sensitive predictor of schizophrenia among children aged 7–8 in the general population; whereas, among high risk children, poor social functioning may be quite sensitive to schizophrenia as early as age 5-6. However, given an estimated effect size (d) of about 1, and an OR of about 5-6, it would be mistaken to try to predict psychosis on this basis. Even with the elevated baseline risk of 15% in the EHRIS, the sensitivity and specificity of such behaviours were too low (109).
Physical examination and neuropsychological test prediction

In the NYHRP, the offspring were tested with neurobehavioral measures at 7-12 years of age and assessed in mid-adulthood for schizophrenia-related diagnoses. Childhood deficits in attention, verbal memory, and gross motor skills identified 83%, 75%, and 58%, respectively, of those with psychoses; 50% were identified by all three variables combined. Encouragingly, the three variables had low deficit rates in the offspring of two other parental groups and were not associated with other psychiatric disorders in any group, but false positive rates were 18-28%, which the authors rightly regarded as insufficient evidence for antipsychotic drug prescribing (111). Michie et al (112), similarly, reported a false positive rate of 21% as unacceptably high in children assessed for sustained attention deficits. Worse, NSS were not predictors of symptoms or schizophrenia in EHRS (113), and cognitive tests were at most weak predictors (114).

Indeed, Pukrop et al (115) recently reviewed 32 relevant cognitive studies and found that investigations of neurocognitive baseline assessments in high-risk samples are inconsistent in terms of the deficits found. Longitudinal studies tend to favour measures of processing speed and of verbal memory and learning as predictive of psychosis, but the weak predictive effects, negative studies and unstable performance argue against the usefulness of cognitive tests in early diagnosis, at least in isolation.

Multivariate prediction

Several studies around the world have now examined the predictive performance of combinations of symptoms and other variables, with mixed results. Even though features like bizarre thinking and schizotypy are commonly replicated, they tend to do so as part of multivariate models which are dissimilar (104,105,107). The North American Prodrome Longitudinal Study (NAPLS) followed up 291 prospectively identified treatment-seeking patients with prodromal syndromes criteria, 35% of whom developed schizophrenia. Of 77 variables, five baseline features contributed uniquely to the prediction of psychosis: a genetic risk for schizophrenia with recent deterioration in functioning (one of the UHR criteria), higher levels of unusual thought content and suspicion/paranoia, greater social impairment, and a history of substance abuse (116). Prediction algorithms combining 2 or 3 of these variables resulted in dramatic increases in positive predictive power (up to 80%) compared with the UHR prodromal criteria alone. The equally impressive European Prediction Of Psychosis (EPOS) study established high inter-rater reliability for the >60 items they examined and optimal prediction with six variables (positive symptoms, bizarre thinking, sleep disturbances, schizotypal PD, highest functioning in the past year, and years of education). This combination gives a positive likelihood ratio above 10 (107). It awaits replication, however, and did not replicate the predictive power of either the Bonn (104) or NAPLS criteria (116).

Blood tests

In theory, the genomic biomarkers described above could potentially predict schizophrenia at an early stage of development, many years before onset. There are, however, only two studies which have taken blood before diagnosis, and both of these in adults rather than children, perhaps for practical and ethical reasons. In the EHRS, NRG1 status was associated with the onset of psychotic symptoms (117), whereas the COMT Val/Met allele polymorphism was the only schizophrenia predictive blood test (Val allele present PPV 39%, NPV 93%; 118). This result gains partial support from replicated work showing a COMT-cannabis interaction (119), although there was no such interaction in the EHRS. Clearly, these results require clarification before genotyping could be employed as a diagnostic marker in high risk groups.

Neuroimaging

There are now a number of studies of people at genetic high risk or with prodromal symptoms who have been imaged at baseline and subsequently examined for transition status, some with follow-up imaging. Reductions of grey matter (GM) density in orbito-frontal cortex (120-122) and medial temporal lobe (120,122) are now clearly replicated in the prodrome to schizophrenia, although the numbers are small. Three studies have taken these analyses further into the clinical domain. Schobel et al (124) found that increases in hippocampal CA1 cerebral blood volumes on contrast enhancement predicted subsequent psychosis with PPV 71% and NPV 82%. Koutsouleris et al (125) found overall SVM classification accuracies of around 90% in discriminating between at risk groups and healthy controls. A receiver operator characteristic curve analysis of GM change in the inferior temporal lobes in the EHRS showed that these were more strongly predictive of schizophrenia than any other variable in that study, with a likelihood ratio of more than 10 (126; PPV 60%, NPV 92%).

It would, of course, be much easier and cheaper to be able to use one baseline scan to predict schizophrenia, and several groups have provided proof of concept studies, although the results are confusing. As Smieskova et al (127) showed in a recent systematic review of the literature, cross-sectional voxel-based morphometry studies have replicated decreased GM in frontal and cingulate cortex in the pre-psychotic, yet whole brain volumes and/or global GM volumes were consistently increased. Indeed, in the EHRS, increased PFC folding on the first scan had a PPV of 67%, our strongest baseline predictor (128). This points to a dramatic reduction in volumes around onset, which could be focus for future investigations, and suggests that analysis techniques which can allow for baseline increases and decreases as well as change may have the best diagnostic performance.
PREDICTING ANTIPSYCHOTIC DRUG TREATMENT RESPONSE

Treatment response is pertinent to clinically relevant pathophysiology to the extent that available treatments address the fundamental disease process or processes rather than being simply ameliorative in some way. We can be confident that antipsychotic drugs treat the hyperdopaminergia associated with positive psychotic symptoms, and even though it is not clear that this is the primary disease process in schizophrenia, there is substantial evidence that this represents a common pathway to acute delusions and hallucinations.

Clinical predictors

Several historical variables have been repeatedly associated with a good response to antipsychotic drugs (including symptom severity, early subjective and objectively rated response to the drug, and the duration of untreated psychosis), but very few researchers have examined their diagnostic properties in prediction (129,130). Recent examples include an attempt to use baseline Positive and Negative Syndrome Scale (PANSS) scores to predict response at week 2, but the predictive values were low (131). Leucht et al (132) have shown that predicting non-response on the Brief Psychiatric Rating Scale (BPRS) at 4 weeks with a PPV of >80% was only possible if there had been absolutely no improvement at all in the first two weeks. The prediction of remission might be improved by the inclusion of 4- and 6-week assessments, but the increase in prediction accuracy is modest at best and unlikely to be clinically useful (133).

The Drug Attitude Inventory is a 30 item self report inventory which has good psychometric properties and diagnostic performance, perhaps because it captures elements of both an early subjective response and positive attitudes to medication (134), which are both associated with compliance. This and standardized symptom severity and outcome ratings might be usefully incorporated into routine clinical practice, at least to help reliably determine people’s attitudes to treatment and whether they have benefitted sufficiently to stay on a treatment.

Biological predictors

Biomarkers of treatment response do not have stiff competition, but they still have a long way to go. Higher antipsychotic drug plasma levels and raised homovanillic acid (HVA) and other peripheral markers in plasma (and CSF) have been repeatedly related to response, but the replicability, diagnostic performance and practicality of this are unclear (135). Further, plasma measures are themselves often at best indirect measure of cortical activity. Most potential pharmacogenetic predictors of antipsychotic drug response have also fallen at the stage of reproducibility. Intriguing findings that the COMT Val allele might predict olanzapine response (136), that the 102-T/C 5-HT2A receptor gene is associated with clozapine response (137), and that the DRD3 Ser allele is associated with poor clozapine response (138) all await external replication. Only the Del allele within the -141C Ins/Del DRD2 polymorphism is consistently associated with (poorer) antipsychotic drug response relative to the Ins/Ins genotype, but even this effect is too small for clinical use (139). The genetics of antipsychotic drug response may therefore be as complicated as the genetics of schizophrenia, and the pharmacogenetics of psychosis might also require multiple gene testing.

Imaging predictors of response

In sharp contrast with the diagnosis and early diagnosis literature reviewed above, structural imaging measures are clearly not associated with treatment response or resistance (140,141). There are, however, quite a number of studies showing that more abnormal computed tomography/sMRI appearances are associated with a generally poor prognosis and a bad outcome. Functional imaging measures show much greater promise, with both reduced basal ganglia metabolism and increased striatal D2 receptor occupancy being repeatedly linked to antipsychotic drug treatment response (135,142).

There is also a strikingly consistent literature on the EEG and treatment response in schizophrenia, in which increased pre- and/or post-treatment alpha-wave EEG activity predicts response to antipsychotics in five out of the six studies we are aware of (143-148). There is enough replication here to justify further studies of PET and EEG of antipsychotic drug response and to begin to evaluate this in terms of their potential clinical significance. Where PET prediction of response could be really useful is in predicting treatment resistance to first or second generation antipsychotics and, even better, response to clozapine, and perhaps also in measuring the response to a single test dose as a means to establish drug and dosage choice for a given patient. Those questions need, however, to be considered in detail by additional studies. The greater availability and lesser cost of EEG make this the most promising potential predictive biomarker of antipsychotic drug response in psychosis for routine clinical use.

CONCLUSIONS AND RESEARCH DIRECTIONS

We have considered the ability of symptoms and signs, and a range of potential biomarkers, as methods of objectively diagnosing schizophrenia in established cases, in predicting transition to psychosis in people at high risk for clinical or genetic reasons, and in predicting treatment response to antipsychotic medication. We have identified what we consider are the best bets for future research evaluation and provided some pointers about how these studies should be conducted and reported (Table 4). Some will say this is all premature. It would certainly be foolish to think that we are ready to employ these measures in clinical practice, but we think that it is long overdue to start considering the variables and methods...
which can take us towards objective diagnostic testing in psychiatry, and to report data in terms of predictive values and likelihood ratios, or at least in such a way as these can be calculated.

The current “gold standard” for the diagnosis of schizophrenia allows reliable diagnoses to be made and communicated, and has some predictive validity in terms of denoting a poor prognosis in most patients. It is frequently stated that these criteria lack biological validity, yet there is no doubt that they have allowed aspects of the pathophysiology of schizophrenia and other psychoses to be elucidated. The symptomatic and biological boundaries between schizophrenia and bipolar disorder may not be discrete (149,150), but where there have been direct comparisons we have been able to highlight some promising leads. We fully acknowledge, and indeed it is a key motivator to writing this article, that our diagnostic gold standard is tarnished and can be variably applied. Replacing our diagnostic gold standard is tarnished – towards objective, etiological and/or pathophysiological measures. We have been overcautious in pursuing this agenda in psychiatry, as a medical discipline, perhaps in part because of the hype and then failure of the dexamethasone suppression test in depression (151).

We regard the diagnosis section of this paper as the most important part, because a suitable patient population is available to all clinicians, a diagnosis is usually already made, and this is therefore where an objective approach would have most impact. Epidemiological risk factors need to be formally evaluated in terms of how much they should rightly increase diagnostic suspicion (or likelihood), especially when considered with other factors, and as potential causal specifiers for psychosis. We also need to determine if there are any objective, reliable “soft signs”, and how these and brief cognitive tests of intellectual decline from premorbid function may perform in clinical practice in terms of their practicality and utility in patients with acute psychosis. Meanwhile, geneticists need to establish how we will know a causal gene when we see one, and how we will manage the patients carrying it. Imaging “biomarkers” perhaps have most promise for diagnostics, but the imaging community needs to develop quantitative techniques that can be applied to individual patients and apply these to the critical distinction between schizophrenia and bipolar disorder with psychotic symptoms. Amygdala volumes may require standardization by age and account for medication if they are to be a distinguishing feature, while dorsolateral PFC activation patterns will require standardization by performance and perhaps IQ, although resting state functional imaging studies may circumvent this.

Making diagnoses at earlier stages in the illness and therapeutic response predictions are not lesser priorities but do seem less practical propositions. Risk factors are all too rare and insufficiently powerful predictors of psychosis to be of great diagnostic value in essentially healthy people, quite apart from the ethical issues inherent in predictive genetic testing and possible prescription of unproven treatments for large numbers of people years before a few become ill. Early diagnosis becomes more practical and ethically straightforward nearer to the time of onset, when the severity of symptoms, thought disruption, schizotypy, cannabis use and brain imaging again look to have promise. It is, however, at least debatable to what extent a predictive test for schizophrenia, or indeed of antipsychotic drug response, would be used, even if predictors were strong, given the limited resources for early intervention services, the restricted choice of treatments currently available, and the lack of availability of imaging and genetic technologies in most clinical settings even in so-called developed nations.

Even more important than the specifics at this stage is the general approach. The one critical aspect of diagnostic studies that is often forgotten is the necessity of a reliable test of the proposed diagnostic aid in a second independent and preferably similarly large cohort, also conducted blind to diagnosis. As fitted models of multiple variables always perform in an “optimistic manner”, or are “over-fitted” on the model-development data, cross-validation in an independent sample is needed to control for tailormade modelling. We are not aware of any examples of this having been done in a truly independent cohort for any of the findings we have described. This requires large scale clinical research studies, which may require support from a variety of informatics approaches, including computational models of the brain/mind, normative and illness databases for comparisons, multivariate prediction algorithms and so on (152,153). Multilevel models including neurobiological, sociobiographical, and environmental variables may increase predictive accuracy, but each additional domain also brings potential variations according to study setting, levels of ex-
posure and inter-rater reliability, as well increasing the risk of false negatives.

The biggest stumbling block clinical researchers may face in trying to set up such studies and change diagnostic practice in psychiatry is concern about how certain one needs to be of an etiological risk factor or pathophysiological mechanism and its specificity before it can be used as a diagnostic aid or test. This is, of course, a legitimate question, but it misses the key point – at least from a clinical perspective – of whether or not the presence of a marker in an individual takes it beyond a threshold where diagnosis or some management strategy which follows from it is likely to be of benefit. Establishing the requisite measures and thresholds will require formal studies in their own right. Clinicians will need to participate in large simple studies to identify the most clinically useful symptoms and signs and tests. This is how medicine works and, with additional study, advances. It is the way psychiatry needs to travel if we are to start to use objective indices to inform psychiatric classification and practice. The future of psychiatry as a medical discipline may depend on it.

References


44. Henry JD, Crawford JR. A meta-analytic review of verbal fluency deficits in schizophrenia relative to other neurocognitive deficits. Cogn Neuropsychiatry 2005;10:1-33.


81. Glahn DC, Ragland JD, Abramoff A et al. Beyond hypofrontality: a quanti-