First-rank symptoms: a first-rank diagnostic test?†
Alex J. Mitchell

SUMMARY
This commentary questions a Cochrane review that examined whether first-rank symptoms are a useful diagnostic tool for differentiating schizophrenia from other psychotic disorders. It concludes that first-rank symptoms are not particularly accurate in this role, although they might be useful initial screening questions in community surveys or waiting-room screening.

DECLARATION OF INTEREST
None

The diagnosis of schizophrenia remains firmly a clinical one based on symptoms and signs, although the approach may be informal, operational or semi-structured. Schizophrenia is typically characterised by at least 6 months of some combination of delusions, hallucinations, and disorganised speech and behaviour causing a deterioration in function. To be more precise about diagnosis, a specific set of symptoms, so-called first-rank symptoms, are often proposed to be the diagnostic gold standard. However, recent work questions the true diagnostic accuracy of these symptoms.

About first-rank symptoms
Kurt Schneider, a German psychiatrist and pupil of Karl Jaspers, proposed that specific symptoms are characteristic of schizophrenia and therefore worthy of ‘first-rank’ status in the diagnostic hierarchy. The English definition of first-rank symptoms arises from the publication of Schneider’s 1946 work Klinische psychopathologie in translation (Schneider 1959) and also from the Present State Examination, a World Health Organization (WHO) questionnaire (Wing 1974). However, neither of these sources is particularly precise in its definitions. Broadly, first-rank symptoms include: auditory hallucinations (including running commentary and voices conversing); somatic hallucinations; thought withdrawal, insertion and interruption; thought broadcasting; delusional perception; and passivity (actions felt to be influenced by external agents). These symptoms have been considered important enough to have been incorporated into several modern diagnostic systems.

Modern diagnosis of schizophrenia
A common approach is described in DSM-IV and DSM-5 (American Psychiatric Association 1994, 2013). DSM-IV lists five key symptom types: (1) delusions; (2) hallucinations; (3) disorganised speech; (4) disorganised or catatonic behaviour; and (5) negative symptoms. Although presence of two of these symptom types is recommended for diagnosis of schizophrenia, diagnosis can be made with just one type if the symptom is auditory hallucinations characterised by running commentary or voices conversing (both first-rank symptoms) or bizarre delusions.

DSM-5 raises the symptom threshold, requiring that an individual exhibit at least two (not one) of the specified symptom types, at least one of which must be one of the first three listed above (delusions, hallucinations, disorganised speech). This means that while previously a single first-rank symptom was sufficient to reach a diagnosis of schizophrenia, first-rank symptoms are now given less weight relative to other symptoms (Shinn 2013).

Soares-Weiser et al’s Cochrane review
My commentary discusses Soares-Weiser et al’s full review (Soares-Weiser 2015), the abstract of which appears in this month’s Cochrane Corner (p. 146, this issue). Box 1 clarifies some terms used.

Specificity and sensitivity of the studies
Early studies were quick to question the specificity of first-rank symptoms when comparing a clinical diagnosis of schizophrenia with that of illnesses such as bipolar disorder (Carpenter 1973). First-rank symptoms are also very common in individuals at high risk of psychosis, but they may or may not predict conversion to a full episode (Morcillo 2015). Thus, it is relevant to consider that the application of a test or tool is dependent not just on the main diagnosis in question but also on the comparator condition. For example, schizophrenia v. aged-matched healthy controls might be relevant to a community field survey. Schizophrenia v. other non-psychotic mental disorder could be useful in primary care or psychiatric practice. Ideally, a test should also be evaluated in all of these situations.

MODERN DIAGNOSIS OF SCHIZOPHRENIA

A common approach is described in DSM-IV and DSM-5 (American Psychiatric Association 1994, 2013). DSM-IV lists five key symptom types: (1) delusions; (2) hallucinations; (3) disorganised speech; (4) disorganised or catatonic behaviour; and (5) negative symptoms. Although presence of two of these symptom types is recommended for diagnosis of schizophrenia, diagnosis can be made with just one type if the symptom is auditory hallucinations characterised by running commentary or voices conversing (both first-rank symptoms) or bizarre delusions.

Soares-Weiser et al’s Cochrane review
My commentary discusses Soares-Weiser et al’s full review (Soares-Weiser 2015), the abstract of which appears in this month’s Cochrane Corner (p. 146, this issue). Box 1 clarifies some terms used.

Specificity and sensitivity of the studies
Early studies were quick to question the specificity of first-rank symptoms when comparing a clinical diagnosis of schizophrenia with that of illnesses such as bipolar disorder (Carpenter 1973). First-rank symptoms are also very common in individuals at high risk of psychosis, but they may or may not predict conversion to a full episode (Morcillo 2015). Thus, it is relevant to consider that the application of a test or tool is dependent not just on the main diagnosis in question but also on the comparator condition. For example, schizophrenia v. aged-matched healthy controls might be relevant to a community field survey. Schizophrenia v. other non-psychotic mental disorder could be useful in primary care or psychiatric practice. Ideally, a test should also be evaluated in all of these situations.

†See p. 146, this issue.
The sensitivity of a test is the proportion of people known to have the disease who test positive: here, the proportion of people with schizophrenia who have first-rank symptoms. A negative test result in conditions with a high-sensitivity test is good at ruling out disease (SnNout). The specificity of a test is the proportion without the disease who test negative for it: here, the proportion without schizophrenia who do not have first-rank symptoms. A positive test result in conditions with a high-specificity test is good at ruling in a disease (SpPin). The positive predictive value is the proportion of people with a positive test result who have the disease: here, the proportion of those with first-rank symptoms who have schizophrenia. The negative predictive value is the proportion of people with a negative test result who do not have the disease: here, the proportion of those without first-rank symptoms who do not have schizophrenia. The likelihood ratio is the likelihood that the same result would be expected in a patient without the disease compared with the likelihood that the same result would be expected in a patient with the disease.

The sensitivity of a test is the proportion of people known to have the disease who test positive: here, the proportion of people with schizophrenia who have first-rank symptoms. A negative test result in conditions with a high-sensitivity test is good at ruling out disease (SnNout). The specificity of a test is the proportion without the disease who test negative for it: here, the proportion without schizophrenia who do not have first-rank symptoms. A positive test result in conditions with a high-specificity test is good at ruling in a disease (SpPin). The positive predictive value is the proportion of people with a positive test result who have the disease: here, the proportion of those with first-rank symptoms who have schizophrenia. The negative predictive value is the proportion of people with a negative test result who do not have the disease: here, the proportion of those without first-rank symptoms who do not have schizophrenia. The likelihood ratio is the likelihood that the same result would be expected in a patient without the disease compared with the likelihood that the same result would be expected in a patient with the disease.

Unfortunately, a significant limitation of Soares-Weiser et al’s review is that the authors excluded case–control studies that involved healthy controls. The prevalence of first-rank symptoms in schizophrenia is one indicator of the value (or sensitivity) of these symptoms as a diagnostic test for the disorder, because such a test should usually be positive in those with the condition. Generally, the prevalence of first-rank symptoms in schizophrenia is reported to range between 25 and 88%, and in this review it was 57%.

The problem of the reference standard

Soares-Weiser et al sought to determine the diagnostic accuracy of one or more first-rank symptoms in the diagnosis of schizophrenia against a diagnosis verified by clinical history and examination by a qualified professional. The sample size was adequate, with 21 studies and a total of 5515 participants included in the analysis. They used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool to rate individual studies. Summary estimates were obtained for sensitivity, specificity and likelihood ratios (other measures of accuracy were omitted).

The main criticism of this review was the significant problem of the reference standard, i.e. diagnosis by history and clinician examination by a qualified professional. In nine studies (43%) the primary authors did not clearly report what methods were used as the reference standard. Four studies assessed patients’ medical records to make a diagnosis – a method that is generally accepted as inaccurate. Furthermore, although operational criteria (e.g. DSM or ICD) were part of the reference standard in all studies (apart from one), a complication is that the reference standard itself also included first-rank symptoms in at least 13, and most likely all, of the studies. This is because primary authors did not systematically exclude use of first-rank symptoms in making the criteria-based diagnosis and this introduces a partial circularity into their studies. Namely, the comparison being made here is one of first-rank symptoms alone v. first-rank symptoms combined with other symptoms in operational criteria.

Related to this there is the significant problem that a diagnosis by a psychiatrist is not necessarily a gold standard. This has been shown in studies in depression and particularly in dementia, where a post-mortem verification of pathology can be obtained. This problem is reduced in part by the use of operational criteria and reduced further by use of semi-structured interviews such as the Structured Clinical Interview for DSM (SCID) and Mini International Neuropsychiatric Interview (MINI).

Other limitations

Soares-Weiser et al made another critical assumption that undermines confidence in their results. In studies that did not specifically use first-rank symptoms to diagnose schizophrenia, but simply measured the prevalence of first-rank symptoms, they assumed that the number of first-rank symptoms reported was the number of first-rank symptoms needed to diagnose schizophrenia. Ideally, this should have been verified on a case by case basis with the original authors.

Regarding masking (blinding), only three studies (14%) reported that the reference standard was interpreted without knowledge of the index test result.

Finally, there was significant heterogeneity in that only 13 studies included only participants with psychosis. Studies in which comorbid psychiatric conditions and related diagnoses such as schizoaffective disorder were combined in the data were not excluded. Indeed, when analysis of results was limited to those with pure schizophrenia alone then summary sensitivity fell to 63.3% (95% CI 56.3–69.9%) and specificity fell to 63.6% (95% CI 48.1–76.7%). Seven studies included all individuals admitted to psychiatric wards with psychotic and non-psychotic symptoms. Six studies included people with first-episode psychosis or first admissions to hospital.

Clinical utility of first-rank symptoms

Ignoring the serious limitations documented above and looking at the results overall, across 20 studies, first-rank symptoms differentiated schizophrenia from all other diagnoses with a sensitivity of 57%
(95% CI 50.4–63.3%) and a specificity of 81.4% (95% CI 74–87.1%). Using the Clinical Utility Index Calculator (www.clinicalutility.co.uk) we can calculate the true value of this test (Table 1). On the basis of a prevalence of 48% reported in Soares-Weiser et al’s review, schizophrenia was unusually common (probably accounted for by the mental health setting of most studies). At this prevalence, the positive predictive value would be 73.9%, meaning that three out of four people with first-rank symptoms would have schizophrenia (the rest would be false positives). The negative predictive value would be 67.2%, meaning that two out three without first-rank symptoms would not have schizophrenia (the remainder would be false negatives).

Given the review result that only 57% of those with schizophrenia have first-rank symptoms (sensitivity) and of those that do, 74% are true positives (positive predictive value), then the overall clinical utility of first-rank symptoms in the confirmatory case-finding diagnosis of schizophrenia (v. all other mental health conditions) can be considered to be ‘poor’. Similarly, given that 81% of those without schizophrenia do not have first-rank symptoms (19% in fact do) (specificity), and of those only 67% are true negatives (negative predictive value), then the clinical utility of first-rank symptoms in ruling out cases of schizophrenia (v. all other mental health conditions) (i.e. screening) can be rated as ‘fair’. Overall, about 7 out of 10 diagnoses based on first-rank symptoms are accurate (that is, they generate true positives or true negatives) and 3 out of 10 are errors. These results, if the methodological problems can be put aside, suggest that first-rank symptoms are surprising poor diagnostically.

Conclusions

A timely and accurate diagnosis of schizophrenia is a priority. Further, there is value in a short series of questions that can be easily applied by mental health professionals, general practitioners and other clinicians. Despite serious methodological limitations, Soares-Weiser et al’s review suggests that, unfortunately, first-rank symptoms do not appear to be particularly accurate in this role. They are present in only about 60% of those with schizophrenia and yet are seen in 20% of those without. However, they may have value in another context. For example, taking the population prevalence of schizophrenia to be about 1% (National Institute of Mental Health 2015), it is likely that first-rank symptoms would be useful as initial screening questions in community surveys or waiting-room screening. That said, a number of studies have documented a continuum of psychotic experiences in about 5% of the general population (van Os 2009). Clearly, more work is required to clarify whether individual first-rank symptoms have particular diagnostic value and whether a combination of symptoms might be more useful.

References


American Psychiatric Association (2013) Diagnostic and Statistical Manual of Mental Disorders (5th edn) (DSM-5). APA.


TABLE 1 Summary of review results

<table>
<thead>
<tr>
<th></th>
<th>Schizophrenia v. other diagnoses</th>
<th>Schizophrenia v. other psychosis</th>
<th>Pure schizophrenia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence, %</td>
<td>48</td>
<td>57</td>
<td>48</td>
</tr>
<tr>
<td>Sensitivity, % (95% CI)</td>
<td>57 (50.4–63.3)</td>
<td>58.0 (50.3–65.3)</td>
<td>63.3</td>
</tr>
<tr>
<td>Specificity, % (95% CI)</td>
<td>81.4 (74–87.1)</td>
<td>74.7 (65.2–82.3)</td>
<td>63.6</td>
</tr>
<tr>
<td>Positive predictive value, %</td>
<td>73.9</td>
<td>75.2</td>
<td>61.6</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>67.2</td>
<td>57.3</td>
<td>65.2</td>
</tr>
<tr>
<td>Clinical utility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive (case-finding)</td>
<td>Poor (0.422)</td>
<td>Poor (0.436)</td>
<td>Poor (0.390)</td>
</tr>
<tr>
<td>Negative (screening)</td>
<td>Fair (0.547)</td>
<td>Poor (0.428)</td>
<td>Poor (0.415)</td>
</tr>
<tr>
<td>Overall accuracy, %</td>
<td>70</td>
<td>65.2</td>
<td>63.5</td>
</tr>
</tbody>
</table>

a. Data analysis carried out using the Clinical Utility Index Calculator (www.clinicalutility.co.uk).