The tail wagging the dog: the diagnostic accuracy of first rank symptoms

Leigh Townsend & **Riccardo De Giorgi**

COMMENTARY ON… COCHRANE CORNER[†]

SUMMARY

Outcomes for people with schizophrenia are improved by expedient diagnosis and specific treatment. ICD-11 and DSM-5 have reduced the importance of Schneider's first rank symptoms (FRS) in the diagnosis of schizophrenia; however, FRS may still offer a useful triage tool for the early identification of schizophrenia and initiation of antipsychotic therapy in high-demand and resource-poor settings. This commentary considers a Cochrane review that assesses the diagnostic accuracy of one or multiple FRS in diagnosing schizophrenia in adults and adolescents.

DECLARATION OF INTEREST

None.

KEYWORDS

Psychopathology; first rank symptoms; diagnostic accuracy.

In the absence of well-validated and distinct biomarkers for schizophrenia, mental health professionals rely on longitudinal psychopathological observation to differentiate schizophrenia from other psychiatric disorders. In 1959, Schneider proposed a set of 'positive' psychotic symptoms, which he termed first rank symptoms (FRS), as distinctive of schizophrenia (Table 1) (Schneider 1959). These were later incorporated into operationalised diagnostic criteria used worldwide in psychiatric practice. In ICD-10 (World Health Organization 1992) and DSM-III and DSM-IV (American Psychiatric Association 1980, 1994), the presence of one FRS was sufficient to make a diagnosis of schizophrenia. Changes in DSM-5 (American Psychiatric Association 2013) and ICD-11 (World Health Organization 2018) have significantly reduced the importance of FRS, removing their special significance in the operational diagnostic threshold. However, FRS still retain their influence, viewed as a crucial part of the psychopathological phenotype of schizophrenia; importantly, they

continue to be taught to new generations of clinicians and to be employed in the assessment of psychiatric patients.

Summary of the Cochrane review

The review by Soares-Weiser *et al* (2015) in this month's Cochrane Corner aimed to assess the diagnostic accuracy of FRS for schizophrenia, compared with assessment by a qualified professional with or without the use of operational criteria and checklists. It included 21 studies reporting the assessment of 6253 adults and adolescents. Results showed that FRS differentiate schizophrenia from all other diagnoses with a sensitivity of 50.4–63.3% and a specificity of 74–87.1%. The authors concluded that FRS are better at 'ruling out' rather than 'ruling in' a diagnosis of schizophrenia and therefore may still be helpful in the initial screening of people with suspected schizophrenia.

Method

MEDLINE, EMBASE and PsycInfo were searched in 2011 and 2012, MEDION in 2013; although the studies included in the review were conducted between 1974 and 2011, about 80% of them dated up to the 1990s. Additional references were identified by hand-searches of the included studies. The review authors included 21 studies evaluating the sensitivity and specificity of FRS (one or multiple) for the diagnosis of schizophrenia compared with a reference standard, irrespective of publication status and language. Both retrospective and prospective studies with consecutive or random participant selection were considered. It should be noted that the majority of these studies were not specifically designed to assess the diagnostic accuracy of FRS.

These studies reported on 6253 participants but only 5515 were included in the review's analysis. Soares-Weiser *et al* do not explicitly state the reason for this; however, participant inclusion criteria were loose, and participants were excluded only if an organic cause of psychosis such as

Leigh Townsend, MBChB, BMedSc (Hons), is a Foundation Year 2 doctor on the Oxford Academic Foundation Programme. Clinically, he currently works as a senior house officer in the Oxford University Hospitals emergency department at the John Radcliffe Hospital, Oxford, UK. As an academic, he works in the Oxford Autoimmune Neurology Group (University of Oxford) on projects aiming to characterise the psychopathological phenotype of NMDA-receptor antibody encephalitis **Riccardo De Giorgi**, MD, MRCPsych, is a Wellcome Trust Doctoral Training Fellow (DPhil in Biomedical and Clinical Sciences) in the Department of Psychiatry at the University of Oxford and an honorary MRCPsych Clinical Fellow with Oxford Health NHS Foundation Trust, Oxford, UK. He works on experimental medicine trials in patients with treatment-resistant depression. **Correspondence** Dr Leigh Townsend, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, Oxford OX3 9DU, UK. Email: leigh.townsend@medsci.ox.ac.uk

[†]See this issue.

TABLE 1 Schneider's first rank symptoms

Symptom	Definition	Example
Auditory hallucination	Auditory perception with no external cause. The particular form is specified: (a) hearing thoughts spoken aloud (b) hearing voices referring to the individual made in the third person (c) taking the form of a commentary	 'I hear my thoughts outside my head' 'The first voice says "He used that fork in an odd way" and then the second replies "Yes, he did"' 'They say "He is sitting down now talking to the psychiatrist"'
Thought withdrawal, insertion, interruption	The individual believes that their thoughts are under the control of an external agency and can be removed, inserted (and perceived to be alien) or interrupted by others	'My thoughts are fine except when Michael Jackson stops them'
Thought broadcasting	The individual believes that their others can hear or are aware of their thoughts	'My thoughts filter out of my head and everyone can pick them up if they walk past'
Somatic hallucination	An hallucination involving the perception of a physical/bodily experience	'I feel them crawling over me'
Delusional perception	A true perception, to which the individual attributes a false meaning	A normal event such as a traffic light turning red may be interpreted as meaning that Martians are about to land
External agency	The individual believes that their actions or feelings are caused/controlled by another person or force	'The CIA controlled my arm'

Modified from Soares-Weiser et al (2015).

infection or alcohol use was highlighted. This may explain why the 738 participants were not included in subsequent analysis.

The index test was the presence of one or multiple FRS. The comparative weight of individual symptoms in diagnosing schizophrenia was not the focus of the review. As there is no gold standard (Box 1) for the diagnosis of schizophrenia, history and clinical examination performed by a qualified professional (e.g. a psychiatrist, nurse or social worker) with or without the use of operational criteria or checklists of symptoms such as ICD-10 and DSM-IV or earlier versions of these criteria, was used as reference standard (Box 1).

The review authors extracted true-positive, true-negative, false-positive and false-negative rates for differentiating schizophrenia from all other diagnoses, from other psychotic diagnoses alone and/or from non-psychotic diagnoses alone. If these data were not available, they attempted to derive them from summary statistics such as sensitivity, specificity (Box 2) and odds ratios, when reported. Meta-analysis including assessment for heterogeneity was performed to derive weighted accuracy summaries (sensitivity and specificity percentages) for distinguishing: schizophrenia from all other diagnoses; schizophrenia from other psychotic disorders; schizophrenia from non-psychotic disorders.

Assessment of methodological quality was made using the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (Whiting 2011), which consists of four domains: patient selection, index test, reference standard, and flow and timing. The quality assessment was not used to exclude studies, but to describe the internal and

BOX 1 Reference standard versus gold standard

A reference standard, when referring to a test for a given condition, is the test against which the test under investigation (the index test) is compared. Ideally, this should be the best test available.

In medicine, the term 'gold standard' is commonly used to describe the test which is most successful at diagnosing a condition within practical and ethical limits. It does not necessarily refer to the test which is most appropriate in all clinical situations, but it should refer to a test that has been experimentally validated and that achieves high sensitivity and specificity. As an example, gadolinium-enhanced magnetic resonance angiography is the gold standard for the diagnosis of aortic dissection, with a sensitivity and specificity of over 95% (Gebker 2007).

Although a perfect test with 100% sensitivity and specificity is not feasible, diagnostic accuracy studies are based on a one-sided comparison of the results of the index test and those of the reference standard. As discrepancies must be assumed to arise from an error in the index test, when a gold-standard test does not exist, limitations in the reference standard may lead to underestimation of the accuracy of the index test.

external validity of the included studies and to make recommendations for the design of future studies.

Results

Twenty-one studies (5079 participants) were included in the meta-analysis assessing the accuracy of FRS in differentiating schizophrenia from all other psychotic and non-psychotic diagnoses; the median sample size was 146 (range 51–1119), the summary sensitivity and specificity were 57.0% (95% CI 50.4–63.3) and 81.4% (95% CI 74.0–87.1) respectively. With regard to using FRS to differentiate schizophrenia from other types of psychosis, 16 studies (4070 participants) with median sample size of 138 (range 30–996) showed summary sensitivity and specificity of 58.0% (95% CI 50.3–65.3) and 74.7% (95% CI 65.2–82.3) respectively. Finally, for using FRS to differentiate schizophrenia from non-psychotic disorders, the meta-analysis included 7 studies (1652 participants) with a median sample size of 134 (range 45–934) and showed summary sensitivity and specificity of 61.8% (95% CI 51.7–71.0) and 94.1% (95% CI 88.0–97.2) respectively.

The investigations of heterogeneity showed no significant difference ($P = 0.1$) in sensitivity and specificity between all patients admitted to a psychiatric ward compared with those with specific psychoses. This is to be expected as the majority of studies were conducted 20–30 years ago and therefore most patients who were admitted probably had a significant disturbance. Importantly, the review

authors were not able to report on the effect of inclusion of FRS within the reference standard, as no studies reported on this factor.

There were several limitations in the quality of the included studies that may have led to an overestimation of test accuracy. Although usable data could be extracted, the majority of the included studies were not designed to assess the diagnostic accuracy (Box 3) of FRS. This meant that methodological details were often poorly reported, the enrolment of participants was not clearly described and participants may have undergone some degree of selection for inclusion in the studies that does not reflect the range of patients that would present in clinical practice. Owing to these limitations the methodological quality of the studies was mostly rated as 'unclear', although the reporting for flow and timing was generally better, with approximately half of the studies rated as at 'low' risk of bias for this parameter.

Discussion

Are FRS characteristic of schizophrenia?

There are two questions that the review by Soares-Weiser *et al* (2015) allows us to probe. The first is a complex, epistemological question as to whether FRS are valid descriptors of the psychiatric syndrome or indeed a physiopathologically distinct disease entity (i.e. 'true schizophrenia'). This is not a trivial problem, because our ability to identify

and differentiate mental illness in order to initiate appropriate treatments depends on a solid, evidence-based description of the disease we intend to observe. Reviews by Nordgaard *et al* (2008), published before the review in question, and by Heinz *et al* (2016), published after it, address this question from multiple perspectives. Both studies emphasise the lack of a robust evidence base to justify the use of FRS as a diagnostic tool for schizophrenia and encourage further work to understand the neurobiology and psychopathology of 'self-disorder' as a marker for schizophrenia. Heinz *et al* argue that the absence of FRS should make a clinician suspicious that an organic or somatic cause might be present (Heinz 2016). They suggest that in situations where an extensive work-up is not feasible, the absence of complex hallucinations and thought disorder described by the FRS may be used to indicate the need for further assessment in an individual with apparent psychiatric features.

Are FRS an accurate screening tool for schizophrenia?

This leads to the second, more pragmatic question: whether FRS is sufficiently accurate to be used as a screening tool for schizophrenia, to triage patients presenting to mental health services. The review authors emphasise the importance of this question, arguing that in low- and middle-income countries (accounting for 70% of the world's population), there is only one psychiatrist for every one million people (McKenzie 2004). Soares-Weiser *et al* identify significant methodological issues and a wide range of sensitivities and specificities in the use of FRS to diagnose schizophrenia. Furthermore, it is highly likely that FRS have been used as part of the reference standard for many of the studies included in the review. The absence of studies specifically designed to answer the question of the diagnostic accuracy of FRS mean that few studies provided the data necessary to attempt to control for this circularity.

Despite the limitations they observed, Soares-Weiser *et al* argue in favour of the use of FRS in regions where there are few psychiatrists per capita and there is a need for simple, effective mental health screening tools to support the professionals delivering the service. They state that FRS perform better at 'ruling out' than 'ruling in' schizophrenia (Soares-Weiser 2015); however, this claim is not supported by the findings of their review. FRS would need to be shown to have a higher sensitivity than specificity in order to be better at 'ruling out' than 'ruling in' a diagnosis of schizophrenia. Disregarding the relative specificity, for FRS to be useful as a rule-out test, a higher sensitivity would

be required: the review's finding of an upper confidence limit of 63.3% for sensitivity in distinguishing schizophrenia from all other diagnoses means that excluding schizophrenia on the basis of the absence of FRS would miss approximately four out of every ten patients with schizophrenia who are assessed.

The specificity of FRS in each of the reported subgroups is higher. Notably, assessing FRS for differentiating schizophrenia from non-psychotic disorders, the meta-analysis of 7 of the 21 studies showed a summary specificity of 94.1% (95% CI 88.0–97.2) compared with the analysis of 16 of the 21 studies to assess FRS for differentiating schizophrenia from other types of psychosis, which showed a summary specificity of 74.7% (95% CI 65.2–82.3). As FRS are descriptions of specific forms of positive symptoms of psychosis, it is arguably unsurprising that applying FRS should differentiate patients expressing these symptoms from non-psychotic patients. The reduction in specificity evidenced when differentiating schizophrenia from other psychotic disorders suggests that focus on the specific modality, form or content of the psychotic features, necessary to apply FRS as a tool, has poor diagnostic value. The inclusion of FRS in the reference standard in the majority of studies, the lack of up-to-date studies, specifically designed to answer the question of FRS' diagnostic accuracy and the evidence of poor reporting of methodology suggest that the data above should not be overinterpreted, as its reliability or generalisability may be limited.

The relevance of the review findings

The impact of this Cochrane review and the weight of similar evidence and expert opinion contributed to alterations in the latest DSM (American Psychiatric Association 2013) and ICD (World Health Organization 2018) criteria for schizophrenia that have de-emphasised FRS in the diagnosis of the disorder. However, this work remains relevant today in highlighting the fundamental problem in the description and diagnosis of psychiatric diseases, the lack of objective and reliable markers around which pathophysiological descriptions, and therefore diagnostic tests, can be constructed (Nordgaard 2008). It is also important to reflect on the historical context in which our current diagnostic framework has evolved. For example, Heinz *et al* argue that Schneider may have emphasised internal experience, requiring patient self-report, over affect or behaviour, requiring the interpretation of another person, in order to avoid observer bias due to the prejudice against the psychiatrically unwell and the significant danger posed to individuals

diagnosed with schizophrenia in Germany at that time (Heinz 2016).

The relegation of FRS in current diagnostic criteria appears justified on the basis of the available evidence; however, this is not to say that psychopathological descriptions should not form part of diagnostic standards. A potential concern is that the lack of studies assessing the diagnostic value of psychopathological features such as FRS may lead to further de-emphasis of psychopathology in the diagnostic criteria. Specifically designed prospective studies, using DSM-5 or ICD-11 criteria as a reference standard, could mitigate the effect of circularity and generate a clearer evidence base for the use of FRS; however, such studies would represent further attempts to compare one diagnostic convention with another without anchoring them around one or more well-validated biological or psychopathological constants (Nordgaard 2008).

Conclusions

This Cochrane review identifies a lack of high-quality evidence for the use of FRS as a diagnostic test for schizophrenia. Soares-Weiser *et al* recommend future research focusing on the utility of FRS as an initial screening test by non-psychiatrists in low-resource settings. Although we do not agree that the evidence supports the use of FRS in this capacity, we agree that better studies are needed, if FRS continues to be employed formally or informally. It is important that future work on diagnosing schizophrenia incorporates a mechanistic understanding of brain function and a self-conscious appreciation of the historical influences which have led to our current understanding of schizophrenia.

References

American Psychiatric Association (1980) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn) (DSM-III). APA.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). APA.

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

Gebker R, Gomaa O, Schnackenburg B, et al (2007) Comparison of different MRI techniques for the assessment of thoracic aortic pathology: 3D contrast enhanced MR angiography, turbo spin echo and balanced steady state free precession. *International Journal of Cardiovascular Imaging*, **23**: 747–56.

Heinz A, Voss M, Lawrie S, et al (2016) Shall we really say goodbye to first rank symptoms? *European Psychiatry*, **37**: 8–13.

McKenzie K, Patel V, Araya R (2004) Learning from low income countries: mental health. *BMJ*, **329**: 1138–40.

Nordgaard J, Sidse M, Arnfred P, et al (2008) The diagnostic status of first-rank symptoms. *Schizophrenia Bulletin*, **34**: 137–54.

Schneider K (1959) *Clinical Psychopathology* (5th edn) (trans MW Hamilton). Grune & Stratton.

Soares-Weiser K, Maayan N, Bergman H, et al (2015) First rank symptoms for schizophrenia. *Cochrane Database of Systematic Reviews*, **1**: CD010653 (doi: 10.1002/14651858.CD010653.pub2).

Whiting PF, Rutjes AWS, Westwood ME, et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, **155**: 529–36.

World Health Organization (1992) *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines.* WHO.

World Health Organization (2018) *ICD-11: International Classification of Diseases 11th Revision.* WHO (https://icd.who.int/en/). Accessed 6 August 2019.