**ROUND THE CORNER**

**The tail wagging the dog - the diagnostic accuracy of first rank symptoms**

**A Commentary on…Cochrane Corner: First Rank Symptoms in Schizophrenia**

https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010653.pub2/full?highlightAbstract=withdrawn%7Crank%7Cs%7Cfirst

**Biography**

Leigh Townsend 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. As an academic, he works within the Oxford Autoimmune Neurology Group (University of Oxford), on projects aiming to characterise the psychopathological phenotype of NMDA-receptor antibody encephalitis.

Riccardo De Giorgi 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. He works on experimental medicine trials in patients with treatment- resistant depression.

**Abstract**

Outcomes for patients with schizophrenia are improved by expedient diagnosis and specific treatment. The 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 month's commentary will consider a Cochrane review by Soares-Weiser *et al.* (2015), which assesses the diagnostic accuracy of one or multiple FRS for diagnosing schizophrenia in adults and adolescents.

**Introduction**

In the absence of well-validated and distinct biomarkers for schizophrenia, mental health professionals rely upon longitudinal psychopathological observation to differentiate schizophrenia from other psychiatric disorders. In 1959, Schneider proposed First Rank symptoms (FRS, i.e. a set of "positive" psychotic symptoms, see Table 1) (**Schneider 1959**) as distinctive of schizophrenia. These were later incorporated into operationalised diagnostic criteria used worldwide in psychiatric practice: in the ICD-10 and DSM-III and -IV, the presence of one FRS was sufficient to make a diagnosis of schizophrenia. The changes in the DSM-5 and ICD-11 have significantly reduced the importance of FRS, removing their special significance in the operational diagnostic threshold. FRS still retain their influence, viewed as a crucial part of the psychopathological phenotype of schizophrenia; importantly, they continue to be taught to the new generations of clinicians and employed in the assessment of psychiatric patients.

Table 1: Schneider's First Rank Symptoms (FRS) - Modified from **Soares-Weiser 2015**

**First Rank Symptoms**

| Symptom | Definition | Example |
|---|---|---|
| **Auditory hallucination** | Auditory perception with no external cause. The particular form is specified:<br>- hearing thoughts spoken aloud<br><br>- hearing voices referring to his/herself made in the third person<br><br>- taking the form of a commentary | "I hear my thoughts outside my head"<br><br>"The first voice says 'He used that fork in an odd way' and then the second replies 'Yes, he did'"<br><br>"They say 'He is sitting down now talking to the psychiatrist.'" |
| **Thought withdrawal, insertion, interruption** | A person's thoughts are under 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** | Others can hear or are aware of the individual's 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 experience with the body | "I feel them crawling over me" |
| **Delusional perception** | A true perception, to which a person attributes a false meaning | A normal event such as a traffic light turning red may be interpreted by the patient as meaning that Martians are about to land |
| **External agency** | The actions or feelings of the individual are caused/controlled by another individual or force | "The CIA controlled my arm" |

**Summary of study**

The Cochrane review by Soares-Weiser et al. (**2015**) aimed to assess the diagnostic accuracy of FRS for schizophrenia, compared to 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 differentiates 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.

**Methods**

Database's searches through MEDLINE, EMBASE, PsycInfo, and MEDION were conducted until December 2013; although the included studies were conducted from 1974 to 2011, about 80% of them dated up to the 90s. Additional references were identified through hand-search 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 state the reason for this explicitly; however, participant inclusion criteria were loose, and participants were only excluded if an organic source for psychosis such as infection or alcohol use was highlighted. This may be the reason that 738 participants were not included in subsequent analysis.

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

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 (see box 2) and odds ratios, when reported. Meta-analysis including assessment for heterogeneity was performed to derive weighted accuracy summaries (sensitivity and specificity %) for distinguishing: schizophrenia from all other diagnosis; schizophrenia from other psychotic disorder; schizophrenia from non-psychotic disorders.

Assessment of methodological quality was made using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool (**Whiting 2011**) consisting 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 external validity of the included studies and to make recommendations for the design of future studies.

**Results**

Twenty-one studies (5 079 participants) were included in the meta-analysis assessing FRS to differentiate schizophrenia from all other psychotic and non-psychotic diagnoses; the median sample size was 146 (range 51 to 1 119), the summary sensitivity and specificity (95% CI) were 57.0% (50.4%

to 63.3%) and 81.4% (74.0% to 87.1%) respectively. With regards to FRS to differentiate schizophrenia from other types of psychosis, 16 studies (4 070 participants) with median sample size of 138 (range 30 to 996) showed summary sensitivity and specificity (95%CI) of 58.0% (50.3% to 65.3%) and 74.7% (65.2% to 82.3%) respectively. Finally, for FRS to differentiate schizophrenia from non-psychotic disorders the meta-analysis consisted of 7 studies (1652 participants) with median sample size of 134 (range 45 to 934) and summary sensitivity and specificity (95% CI) of 61.8% (51.7% to 71.0%) and 94.1% (88.0% to 97.2%) respectively.

The investigations of heterogeneity showed no significant difference (p=0.1) in sensitivity and specificity between all admissions to a psychiatric ward compared to those with specific psychoses. This is to be expected as the majority of studies were conducted 20 to 30 years ago and therefore most patients who were hospitalised likely had a significant disturbance. Importantly, the 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 the overestimation of test accuracy. Though useable data could be extracted, the majority of the included studies were not designed to assess the diagnostic test accuracy (see Box 3) of FRS. This meant that methodological details were often poorly reported, the enrolment of participants was not clearly stated, and participants may have undergone some degree of selection to be included in the studies that does not reflect the range of patients that would present in clinical practice. The methodological quality of the studies was mostly rated as "unclear" due to these limitations, although the reporting for flow and timing was generally better with approximately half the studies rated as at "low" risk of bias for this parameter.


**Discussion**

There are two questions which the review by Soares-Weiser et al. (**2015**) can 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**), prior to the review in question, and by Heinz et al. (**2016**) afterwards, 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 for the presence of an organic or somatic cause (**Heinz 2016**). They suggest that in situations where an extensive work-up may not be feasible, the absence of complex hallucinations and thought disorder described by the FRS may indicate the need for further assessment in an individual with apparent psychiatric features.

This leads onto 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 issue, arguing that in low- and middle-income countries (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 these studies. The absence of studies specifically designed to answer the question of FRS' accuracy 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 places where there are far fewer 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 performs better at 'ruling out' than 'ruling in' schizophrenia (**Soares-Weiser 2015**); however, this claim is not supported by the findings of the 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: estimating based on the review's results, a sensitivity of 63.3% to distinguish schizophrenia from all other diagnoses, means that excluding schizophrenia based on the absence of FRS would miss approximately four out of every ten patients with schizophrenia that 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% (88.0% to 97.2%) compared to 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% (65.2% to 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 impact of this Cochrane review and the weight of similar evidence and expert opinion prior to the publication of the latest DSM and ICD criteria for schizophrenia contributed to alterations which have de-emphasised FRS in the diagnosis of schizophrenia. However, this work remains relevant today in highlighting the fundamental issue 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 upon the historical context through 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 based on 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, these studies represent further attempts to compare one diagnostic convention to another without anchoring them around one or more well-validated biological or psychopathological constants (**Nordgaard 2008**).

**Conclusion**

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. While 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.

**Boxes**

Box 1 - gold standard versus reference 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 and in medicine, the term '**gold standard**', is commonly used to describe the test which is most successful at diagnosing the 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 which 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**).

While a perfect test which is 100% sensitive and specific is not feasibly obtainable, diagnostic accuracy studies are based upon the one-sided comparison between the results of the index test and those of 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.

Box 2 - sensitivity and specificity:
The **sensitivity** of a test refers to the proportion of patients with the disease in question, who are identified as having the disease, by their test result. Mathematically:

*true positives / (true positives + false negatives)*

When a test has a high sensitivity, a negative test result can be useful for 'ruling out' the disease as it is unlikely to occur if the disease is present.

The **specificity** of a test refers to the proportion of patients, who do not have the disease in question, who are identified as being disease-free by their test result. Mathematically:

*true negative / (true negative + false positive)*

When a test has a high specificity, a positive test result tells you that the disease is likely to be present. In other words, it 'rules in' the disease.

Useful acronym – *SpIn, rule in; SnOut, rule out*

Box 3 - accuracy:
In mathematical terms, the accuracy of a diagnostic test is defined as:

*(true negative + true positive) / (true negative + true positive + false negative + false positive)*

In other words:

*number of correct assessments / total number of assessments*

The accuracy of a test represents the proportion results which will be true, both true positive and true negative, thus measuring the reliability of a diagnostic test for a specific disease. The accuracy of a diagnostic test can also be calculated from a test's sensitivity and specificity and a disease's prevalence (if known) according to the formula:

*((sensitivity) x (prevalence)) + ((specificity) x (1 - prevalence))*

This means that, even if a test has high sensitivity and specificity, the overall test accuracy will be low if the disease in question is rare.

**References**

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. *Int J Cardiovasc Imaging*, **23** (6): 747.

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

McKenzie K, Patel V, Araya R. (2004) Learning from low income countries: mental health. *British Medical Journal*, **329** (7475): 1138–40.

Nordgaard J, Sidse, M. Arnfred, P. *et al*. (2008) The Diagnostic Status of First-Rank Symptoms. *Schizophrenia Bulletin*, **34** (1) : 137–154.

Schneider, K. (1959). *Clinical psychopathology (Trans by M W Hamilton) (5th ed.)*. Oxford, England: 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** (8):529–36.