What are outcome assessment tools?

Standardised assessment tools are vital for assessing a range of variables including symptoms, functioning and quality of life. The quality of these tools can be measured in various ways. ‘Reliability’ refers to the reproducibility of an instrument’s results across different assessors, settings and times. ‘Construct validity’ is the extent to which an instrument measures the theoretical construct it was designed to measure. This involves ‘convergent validity’ (the degree of correlation between different scales measuring the same construct); and ‘divergent validity’ (the lack of correlation between scales measuring different constructs). ‘Known groups’ validity’ is the extent to which an instrument can demonstrate different scores for groups known to vary on the variables being measured. ‘Content validity’ is the extent to which each item on a scale represents the construct being measured, and ‘internal consistency’ is the degree of correlation between items within a scale. ‘Predictive validity’ refers to sensitivity, which is the proportion of correctly identified positives (e.g. high ratings on a scale), and specificity, which is the proportion of correctly identified negatives (e.g. low ratings on a scale). ‘Responsiveness’ is the extent to which an instrument can detect clinically significant or practically important changes over time.

What is the evidence for outcome assessment tools?

Moderate quality evidence suggests the Brief Psychiatric Rating Scale can be factored into five discrete components, comprising positive, negative, and affective symptoms, resistance (hostility) and activation (excitement). The Positive and Negative Syndrome Scale showed a similar factor structure but included a larger number of items in the negative symptom factor and enough items for a discrete disorganisation factor.

Moderate to high quality evidence found good predictive value of the Brief Psychiatric Rating Scale and the Positive and Negative Syndrome Scale for predicting non-response to antipsychotic treatment at 4-12 weeks. There is good predictive validity of the Historical, Clinical and Risk Management-20 scale for predicting aggression in psychiatric facilities. Moderate to low quality evidence suggests the McNiel-Binder Violence Screening Checklist, and the Brøset Violence Checklist may also be effective for predicting aggression or violence, however, the Violence Risk Appraisal Guide had poor predictive validity in people with schizophrenia living in the community.

Moderate quality evidence suggests good inter-rater reliability and some predictive validity for tools assessing duration of untreated psychosis, psychosis onset and treatment onset. Moderate to low quality evidence suggests the Recovery Assessment Scale has the best psychometric properties for measuring personal recovery; it has good construct validity, content validity, internal consistency, test-retest reliability, administrator-friendliness, and has been translated to languages other than English. However, its user-friendly rating is poor.

Other scales rating personal recovery with reasonable psychometric properties include; the Self-Identified Stage of Recovery scale, which has good construct and content validity, good internal consistency but poor test-retest reliability, and good user-friendliness, and has also been translated to languages other than English. The Mental Health Recovery Measure has good content validity but poor construct validity, and good internal consistency and test-retest validity. Moderate to low quality evidence also suggests that reliability is good for the following instruments assessing depressive symptoms: Brief Psychiatric Rating Scale-Depression, Positive and Negative Syndrome scale-Depression, Hamilton Rating Scale for Depression, Montgomery Asberg Depression Rating Scale, Calgary Depression Scale for Schizophrenia, and Beck Depression Inventory. The best concurrent validity indices are reported for the Calgary Depression Scale for Schizophrenia, and the Montgomery Asberg Depression Rating Scale. The highest ranges for sensitivity and specificity were reported for the Calgary Depression Scale for Schizophrenia. Moderate quality evidence suggests good ‘known groups’ validity for the Short Form health survey-36, but inconsistent convergent validity and poor responsiveness.

For further information see the technical table

HOW YOUR SUPPORT HELPS

We are able to make significant advances due to the generosity of countless people. Your donation allows us to continue to work towards transforming lives. For information on how you can support our research, phone 1300 888 019 or make a secure donation at neura.edu.au/donate/schizophrenia.