Medical Research Archives

European Society of Medicine

OPEN ACCESS

Published: February 29, 2024

Citation: Echevarria B, Klein H, et al., 2024. Measuring Psychosis: A Review of Widely Used Clinical Outcome Assessments in Clinical Trials, Medical Research Archives, [online] 12(2). https://doi.org/10.18103/mra.v12i2.4884

DOI
https://doi.org/10.18103/mra.v12i2.4884

ISSN: 2375-1924

REVIEW ARTICLE

Measuring Psychosis: A Review of Widely Used Clinical Outcome Assessments in Clinical Trials

Barbara Echevarria, PhD[1], Hans Klein, PhD[1], Mark G. Opler, PhD, MPH*[1]

[1]WCG Clinical, Inc. Princeton, NJ, USA

ORCiD
Barbara Echevarria: 0009-0006-6445-6404
Mark G. Opler: 0009-0003-2205-7630
Hans Klein: 0000-0001-7715-8154

***Corresponding author:** mgo4@caa.columbia.edu

ABSTRACT

Accurate diagnosis and assessment of symptoms is essential to the appropriate treatment and management of individuals with psychotic disorders. It also plays a pivotal role in research, ensuring that enrolled subjects meet study-specific inclusion requirements and that their progress is accurately tracked throughout the study. Clinical outcome assessments are standardized instruments designed specifically to measure symptoms and their impact on a patient's life and functioning. These instruments can be used to diagnose, categorize, and track symptom severity, and to measure the functional impact of disease on a patient's quality of life. Clinical outcome assessments play a pivotal role in advancing our understanding and treatment of schizophrenia and other psychotic disorders, providing the ability to evaluate the safety and efficacy of emerging treatments and interventions. In this review article we provide an overview of the most widely used clinical outcome assessments in schizophrenia and other psychotic disorders research with a focus on clinical trials, and delve into the specific instruments commonly used to diagnose, measure symptom severity, and assess cognitive ability. A brief description of each instrument and its role in clinical outcome assessment is provided, along with advantages and limitations in implementation. Additionally, the clinician's perspective on the administration and scoring of these scales is included, where relevant. The goal is to familiarize new clinicians and researchers with the available assessment tools, highlighting the advantages, limitations, and any other relevant information that can aid in the selection of the appropriate measurement tools for their patients and studies. Finally, we briefly discuss our view on the future of clinical outcome assessments in schizophrenia and other psychotic disorders research and clinical trials.

Introduction

Psychotic disorders are characterized by variable phenotypic expression and a complex etiology. Accurate diagnosis and assessment of symptoms is essential to the appropriate treatment and management of individuals with psychotic disorders; it also plays a pivotal role in research, ensuring that enrolled subjects meet study-specific inclusion requirements and that their progress is accurately tracked throughout the study. Despite the importance of accurate measurement, and due to the heterogenous and diverse nature of the symptoms associated with these disorders, assessment and diagnosis of schizophrenia and other psychotic disorders remains a challenging task for many clinicians and researchers. Clinical outcome assessments (COAs) are standardized instruments designed to measure symptoms and their impact on a patient's life and functioning. They are commonly used in research to assess eligibility for and efficacy of treatment interventions and to track disease progression, and take the form of questionnaires, interviews, physical examinations, and rating scales. Accurate implementation depends on the subjective judgment of the patient and/or clinician or that of an observer or informant.[1] When administered appropriately, they allow for precise and reliable diagnosis and symptom tracking from the beginning of the study through completion. They also play a pivotal role in advancing our understanding and treatment of schizophrenia and other psychotic disorders, providing the ability to evaluate the safety and efficacy of emerging treatments and interventions. For this reason, selection of appropriate COAs is key to the success of any research study or clinical trial.

In this review article, we aim to provide an overview of the most widely used COAs in schizophrenia and other psychotic disorders research, with a focus on clinical trials. We will delve into the specific COAs commonly used to diagnose, measure symptom severity, and assess cognitive ability, and provide a brief description of each instrument, along with advantages and limitations in implementation. Additionally, the clinician's perspective on the administration and scoring of these scales is included, where relevant. The goal of this review is to familiarize new clinicians and researchers with the available tools to measure psychotic symptoms, highlighting the advantages, limitations, and any other relevant information that can aid in the selection of the appropriate measurement tools for their patients and studies. Finally, we will briefly discuss our view on the future of COAs in schizophrenia and other psychotic disorders research and clinical trials.

Diagnostic tools for schizophrenia and other psychotic disorders

Diagnosing schizophrenia, and assessing symptoms reliably, remains the responsibility of mental health professionals. While strides have been made to identify potential biomarkers and other objective techniques (e.g., neuroimaging) to aid in diagnosis, no methods have been validated to date that can reliably diagnose beyond symptom characterization. Diagnosis in research poses a complex set of challenges for clinicians and researchers due to the intricate nature of the disorder. Schizophrenia is characterized by a range of symptoms that can vary widely from one individual to another, and that may overlap with other psychiatric conditions. Some symptoms, such as delusions and hallucinations, may be present in other psychiatric disorders, like major depressive disorder or bipolar disorder. This overlap can lead to misdiagnosis if the proper diagnostic tools are not used. Application of appropriate differential diagnosis methodologies is key to ensuring an accurate classification, and can ensure that subjects are correctly enrolled in clinical trials.

There are two widely-used instruments that research studies employ in attempts to reliably diagnose schizophrenia and rule out potential exclusionary diagnoses, the Mini International Neuropsychiatric Interview (MINI),[2] and the Structured Clinical Interview for DSM-5 (SCID-5).[3]

MINI INTERNATIONAL NEUROPSYCHIATRIC INTERVIEW

MINI is a structured diagnostic interview designed to provide a brief but comprehensive evaluation of a patient's mental health symptoms and diagnoses. It is not specific to schizophrenia, consisting of a series of modules, each dedicated to a specific disorder or diagnostic category and covering a wide range of conditions including mood, anxiety, substance use, and psychotic disorders.

MINI for Psychotic Disorders, a version of the MINI designed specifically for diagnosing psychotic disorders, includes an expanded module (K. Psychotic Disorders and Mood Disorders with Psychotic Features) providing a more detailed set of questions for each of the nine psychotic disorders than those included in the standard MINI. This additional module is intended for clinical and research settings where psychotic disorders are a focus of interest and where a rigorous differential diagnosis (e.g., schizophrenia vs schizoaffective disorder) is required.

MINI has become the most accepted diagnostic/screening tool in clinical trials in recent years, and is translated and validated for use in around 70 languages. During the interview, the clinician asks a standardized set of "yes/no" questions for each module. If a symptom is reported present, the clinician then asks additional questions to further establish the presence and severity of that symptom. Scoring MINI involves a systematic evaluation of the patient's responses, and in some cases clinician observations, to determine whether the criteria for each disorder assessed is met. The diagnostic criteria for each disorder are based on the two widely known psychiatric classification manuals, The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)[4] and International Classification of Diseases, Tenth Revision (ICD-10).[5] MINI also provides guidance for the classification of symptom severity and conducting differential diagnosis via outlined algorithms.

Advantages

MINI is relatively brief to administer for an experienced clinician. The estimated time to complete has been reported at 15 minutes,[2] although this may not be the standard experience when assessing unknown patients, or for patients that endorse several, disparate symptoms (as is typical in schizophrenia spectrum disorders). Because of its efficiency and brevity, it has become the preferred tool in clinical trials and research studies, where designated diagnosticians have limited time to conduct interviews, or where study screening visits may be already long and burdensome for the patient.

Because it offers a structured and standardized format, MINI can help ensure that clinicians across sites are assessing diagnostic information systematically and consistently for each study participant.

Furthermore, this assessment can be implemented for use in electronic COA systems, providing an advantage to clinicians, with differential diagnostic algorithms implemented directly within the form and reducing greatly the potential for diagnostic error.

Challenges

MINI, because of its brevity, is not as comprehensive as a thorough diagnostic assessment, and is generally used as a screening or confirmatory tool, rather than for a full exploration of a participant's psychiatric history. It may require additional review of collateral sources of information to confirm the accuracy of the diagnosis. This is particularly relevant in psychotic disorders, where patients tend to not be the best reporters due to the nature of their symptoms.

MINI for Psychotic Disorders requires clinicians to perform complex diagnostic decisions following outlined algorithms. This can lead to diagnostic errors and inadvertently arriving at incorrect but related diagnosis with psychotic and mood symptoms.[6]

While MINI covers the most common psychiatric disorders, the scope of the instrument is limited, and may not address less prevalent symptoms and conditions.

Clinician Perspective

MINI can be a quick and rater-friendly diagnostic tool when used to confirm the diagnosis of an already well-known patient. When administered to a new or unknown subject in a clinical trial the structured questions do not provide sufficient information to confidently rate many items, requiring the clinician to come up with additional follow-up questions of their own. When administered on paper, the differential diagnosis algorithms for psychotic and mood disorders can be very difficult to navigate and prone to diagnostic error. Use of the electronic version of MINI alleviates some of the burden in selecting the correct algorithms.

STRUCTURED CLINICAL INTERVIEW FOR DSM-5

SCID-5 is a comprehensive semi-structured interview guide used to accurately diagnose common DSM-5 disorders. SCID-5 is used both in clinical practice and research studies for making accurate psychiatric diagnoses. It is designed to be conducted by experienced clinicians or trained mental-health professionals who are familiar with DSM-5 criteria. SCID-5 can be used to interview individuals with or without known psychiatric history.

SCID-5 offers a standardized format comprised of different modules, which make it easy to adapt and tailor to specific study needs. It usually starts with an overview module, where the clinician collects the psychiatric history of the patient. The interview then guides the clinician through the series of modules that have been selected for the study, which comprise a set of questions carefully aligned with DSM-5 diagnostic criteria. Researchers can choose modules based on their specific study objectives.

SCID-5 also comes with a comprehensive user guide, guiding the clinician step-by-step through the instrument and diagnostic process.

There are currently five available versions of SCID-5:

- Structured Clinical Interview for DSM-5 Research Version (SCID-5-RV):

The most comprehensive version of SCID-5. It includes all relevant subtypes and severities, and course specifiers for each diagnosis. It is highly customizable to meet the study objectives

- Structured Clinical Interview for DSM-5 Clinical Trials Version (SCID-5-CT): Adaptation of SCID-5-RV for use in clinical trials. It offers a more streamlined approach to assessing typical inclusion/exclusion criteria in clinical trials. Usually, it is customized for each study protocol
- Structured Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD): Used to evaluate the ten personality disorders included in DSM-5
- Structured Clinical Interview for DSM-5, Clinician Version (SCID-5-CV): Clinician version. Covers the most common psychiatric disorders seen in clinical settings[7,8]
- Quick Structured Clinical Interview for DSM-5 (Quick-SCID):
  The newest version. A briefer and fully structured assessment which can be completed in under 30 minutes. It is similar to MINI, with questions requiring only "Yes/No" answers.

Advantages

SCID-5 is often considered the "gold standard" diagnostic tool in psychiatric research. Diagnostic questions are tailored to DSM-5 criteria. It allows for comprehensive and in-depth assessment of symptoms, including collection of psychiatric history, leading to a deeper understanding of the subject's condition.

Disadvantages

Because of its semi-structured nature, it may require more clinical experience to administer, as clinicians are required to use clinical judgment to guide their questions and make diagnostic decisions. It takes significantly more time than MINI to administer. The combination of time to administer, plus the requirement of significant clinical expertise, makes it less practical in certain settings, where resources and time may be limited. Additionally, because of its highly customized approach, is not a good candidate for deployment as an electronic COA.

Clinician Perspective

SCID-5 resembles an in-depth clinical interview and provides a much deeper understanding of the patient's history of symptoms vs MINI. It is optimally used when clinicians are evaluating a patient for the first time, and have time to do a thorough assessment. SCID-5 can take up to one hour to administer.

In summary, the choice between MINI and SCID-5 as a schizophrenia diagnostic depends on the aims/goals of the study and the resources available. The brevity of MINI makes it suitable for studies where a quick screening assessment is needed to confirm the subject's diagnosis, whereas SCID-5 is preferred in studies where a more detailed and accurate understanding of a patient's symptoms and diagnosis is required, and where experienced clinicians are available.

Symptom severity clinical outcome assessments in schizophrenia and other psychotic disorders

Symptom severity COAs aim to evaluate the presence and severity of psychopathology symptoms and overall subject status. These COAs are essential in evaluating treatment effectiveness and disease progression. Several outcome assessments used to assess symptom severity in the context of research and clinical trials are reviewed below, including recently developed scales for the assessment of negative symptoms.

POSITIVE AND NEGATIVE SYNDROME SCALE

The Positive and Negative Syndrome Scale (PANSS) has been the most widely used tool for assessing the severity of symptoms in schizophrenia since it was first introduced in 1987.[9] It can be applied in both clinical and research settings. PANSS uses a quantitative approach to symptoms assessment. It consists of 30 items, each rated on a 1–7 scale, where 1 means the symptom is absent and 7 present at the most extreme severity. The individual scores for each or the 30 items are also used to produce a total score. PANSS total score serves as the primary endpoint to measure change associated with a given intervention in most clinical trials. To examine the potential impact on certain symptom domains, the 30 items can be divided into three subscales totals (positive symptoms, negative symptoms, and general psychopathology); researchers have also derived several alternative factor scores that can be used as endpoints.[10,11]

Each PANSS item is rated with the aid of a reference manual that provides a detailed description of the symptom that each item is rating, specifies the basis for rating (e.g., patient report, informant report, or both), and provides anchor benchmarks for the frequency and severity thresholds for each rating. To standardize administration, the Structured Clinical Interview for PANSS (SCI-PANSS), a semi-structured interview guide, can be used to collect necessary information to rate items based on subject report; an analogous informant questionnaire (IQ-PANSS) can be used to

collect information from an informant or caregiver.[12]

Advantages of PANSS

PANSS is well established and widely accepted as a measure for assessing presence and severity of psychotic symptoms in clinical trials. It is a comprehensive scale, assessing a wide range of symptoms associated with schizophrenia and other psychotic disorders. It has been translated and validated in many languages, and its widespread use makes it a valuable tool for comparing research findings across studies involving patients in diverse cultural environments.[13] PANSS has demonstrated good reliability and validity,[14,15] and can yield consistent results when administered by different clinicians.[16,17] Its semi-structured interview format allows clinicians and raters to collect information in a standardized manner by following a script, while allowing flexibility to ask additional questions and clarify inconsistent or ambiguous information. The quantitative approach to scoring, including precise anchor descriptions for each rating, ensures that raters adhere to similar scoring conventions when rating, thus ensuring consistency in severity interpretation.

Weaknesses and Critiques of PANSS

PANSS is a relatively lengthy assessment, taking on average 30-45 minutes to administer and rate. It requires additional caregiver or informant interviews, which in addition to adding burden, may limit the validity of the assessment if the subject does not have an involved caregiver.[18] PANSS ratings are largely dependent on the administering clinician's judgment and the accuracy of informant reports. This can introduce subjectivity into the scoring process. SCI-PANSS can be difficult to navigate for novice raters, who may require extensive training and practice to master the dynamic administration and adhere to its complex skip logic based on subject responses.[19] Some items may not be clearly defined or easily interpreted by clinicians. Quality and psychometric properties of translated versions of PANSS can vary, and researchers and clinicians should consider cultural and linguistic nuances to ensure that the assessment reflect symptoms accurately within the patient's cultural context.

Finally, PANSS has been criticized for placing significant emphasis on the evaluation of positive and negative symptoms, potentially overlooking other important dimensions of schizophrenia, such as cognitive deficits, social cognition, or other functional deficits frequently associated with the disease.

Clinician Perspective

PANSS can be extremely challenging to administer for new raters, although administration significantly improves with practice. Both administration and rating require a significant depth of understanding of psychotic disorders and the full spectrum of symptoms: novice raters can fall short.

SIX-ITEM POSITIVE AND NEGATIVE SYNDROME SCALE

The six-item Positive And Negative Syndrome Scale (PANSS-6) is a shortened version of PANSS, designed to provide a quicker assessment of overall symptom severity by focusing on a subset of key positive (P) and negative (N) symptoms from PANSS. PANSS-6 consists of the following items: P1-Delusions, P2-Conceptual disorganization, P3-Hallucinations, N1-Blunted Affect, N4-Social withdrawal, N6-Lack of spontaneity and flow of conversation.[20] The Simplified Negative and Positive Symptoms Interview (SNAPSI) is a valid interview to collect the required information to accurately rate all 6 items in addition to allowing the rater to score a variety of other scales, such as the Clinical Global Impression (CGI) scale, Brief Psychiatric Rating Scale (BPRS), and 4-Item Negative Symptom Assessment (NSA-4).[21,22] PANSS-6 can be a valuable tool in research and clinical settings where a briefer and more streamlined assessment is desired to save time and reduce patient and clinician burden. PANSS-6 has been validated against PANSS, and been shown to be as effective at measuring treatment effect.[23] However, it is important to note that PANSS-6 does not capture the full spectrum of symptoms, as assessed in the full PANSS, and should be considered only when the concern in change in positive and negative symptoms are primary. Use of PANSS-6 in prospective clinical trials is still required.

BRIEF PSYCHIATRIC RATING SCALE

The Brief Psychiatric Rating Scale (BPRS) assesses overall psychiatric symptomatology.[24] The original scale, published in 1962, consisted of 16 items, which was later expanded to 18 items,[25] and more recently to 24 items.[26] BPRS has been used to assess symptoms associated with psychotic disorders, but its use can be extended to evaluate other psychiatric disorders and conditions.[27,28] Information to rate according to BPRS is collected through a relatively brief clinical interview with the subject.[29] Each of the items is scored from 1–7, where 1 means the symptom is absent and 7 present at the most extreme severity. It has been used extensively to track changes in overall psychopathology over time. BPRS yields several

factor scores in addition to a total score. BPRS served as the basis for the creation of PANSS.[30]

Advantages

BPRS is relatively brief to administer (under 30 minutes), and provides a thorough assessment of psychotic symptoms and general psychopathology, making it suitable for use in cross-diagnostic trials that include disorders beyond schizophrenia spectrum disorders (e.g., mood disorders with psychotic features). The 24-item version is preferred over the older versions due to the expanded symptom constructs examined; the 24-item version also comes with a detailed administration and scoring manual.

Disadvantages and Critique

Similar to PANSS, BPRS requires clinicians to have significant knowledge of psychiatric symptomatology and extensive training on administration and scoring of the scale. As BPRS focuses on psychotic features, the scale does not provide a thorough assessment of negative symptoms.

Clinician Perspective

The 24-item version of BPRS is relatively brief and easy to administer and rate (with some training).

SCALE FOR THE ASSESSMENT OF POSITIVE SYMPTOMS AND SCALE FOR THE ASSESSMENT OF NEGATIVE SYMPTOMS

The Scale for the Assessment of Positive Symptoms (SAPS)[31] and the Scale for the Assessment of Negative Symptoms (SANS)[32] are two separate scales used to assess specific positive and negative symptoms in schizophrenia spectrum disorders. SANS was the first COA specifically developed to assess negative symptoms; it includes the assessment of five dimensions of negative symptoms: alogia, affective blunting, avolition-apathy, anhedonia-asociality, and attentional impairment. SAPS focuses on the assessment of positive symptoms, consisting of 34 items across four dimensions: hallucinations, delusions, bizarre behavior, and positive formal thought disorder. SAPS has been adopted for the assessment of psychotic symptoms in Parkinson's disease.[33-35]

Strengths and Advantages

Each scale provides a separate and in-depth assessment of positive and negative symptoms associated with schizophrenia.

Disadvantages and Critique

The validity and reliability of SANS and SAPS have been questioned, and these scales are not currently widely used in clinical trials. SANS has been criticized for its inclusion of items that do not belong with negative symptoms,[36] and maintaining separate scales for positive and negative symptoms potentially reinforces the outdated dualistic view of schizophrenia.[37]

Assessments of negative symptoms

Assessing negative symptoms is a critical component of evaluating patients with schizophrenia and other psychotic disorders. These symptoms refer to deficits or reductions in normal emotional and behavioral functioning. Measurement of negative symptoms can be challenging for clinicians, as subjects with schizophrenia usually lack insight into these symptoms, reducing the likelihood of reporting them. Clinicians must pay attention to the overall symptomology of the patient, particularly to signs that may indicate the presence of negative symptoms such as affective flattening, lack of social engagement, reduced motivation, etc. In addition to SANS and PANSS, which can be used to measure negative symptoms, newer rating scales have been developed in the last 2–3 decades to systematically measure and track severity of negative symptoms over time; these include the Negative Symptom Assessment (NSA-16), the Clinical Assessment Interview for Negative Symptoms (CAINS), and the Brief Negative Symptom Scale (BNSS).

16-ITEM NEGATIVE SYMPTOM ASSESSMENT

Initially developed in 1989, the 16-Item Negative Symptom Assessment (NSA-16) is a 16-item clinician-rated scale measuring the presence, severity, and range of negative symptoms across five domains: communication, emotion/affect, social involvement, motivation, and retardation.[38] Information collected through a semi-structured interview with the subject is used to rate the 16 items. The NSA-16 manual provides a detailed scoring guide for each item with precise severity anchor descriptions, and also for rating global negative symptom severity. Symptom severity ratings are established as compared to a healthy individual in their twenties. NSA-16 focuses on rating behavior, not psychopathology.

Strengths and Advantages

NSA-16 is a valid scale for measuring negative symptoms[39]; it has been shown to be sensitive to change,[40] and correlates with measures of functioning.[41,42] NSA-16 has been shown to have good interrater reliability when administered across different cultures and languages.[43]

Disadvantages and Critique

NSA-16 has been criticized for focusing exclusively on behaviors, rather that patient's inner experience (e.g., lack of ability to experience pleasure, lack of

interest and motivation),[44] and, as such, studies targeting negative symptoms reduction have adopted newer scales following US National Institute of Mental Health-Measurement and Treatment Research to Improve Cognition in Schizophrenia (NIMH-MATRICS) consensus statement recommendations.[36]

Clinician Experience

The NSA-16 semi-structured interview guide is fairly easy to administer as it focuses exclusively on a subject's behavior, regardless of etiology. The most challenging aspect of NSA-16 is related to scoring each item in comparison to a healthy individual in their twenties, which can introduce bias as interpretation from one clinician to another can differ depending on the clinician's cultural and socioeconomic background.

4-ITEM NEGATIVE SYMPTOM ASSESSMENT

The 4-Item Negative Symptom Assessment (NSA-4) is a shortened version of NSA-16, comprising only four items: restricted speech quantity, reduced emotion, reduced social drive, and reduced interests.[45] It allows for a quick screening assessment of negative symptoms. In addition to producing scores for the four items, clinicians can derive a global rating of the overall impression of negative symptom severity. NSA-4 has been found to be comparable in accuracy to the full-length NSA-16, and to correlate well with other measures of negative symptoms[46]; this makes it a valuable tool for quickly assessing negative symptoms in clinical settings, where time limitations exist. As with the full-length version, NSA-4 focuses on rating behavior, not psychopathology, and thus does not capture patients' inner experiences.

CLINICAL ASSESSMENT INTERVIEW FOR NEGATIVE SYMPTOMS AND BRIEF NEGATIVE SYMPTOM ASSESSMENT

The Clinical Assessment Interview for Negative Symptoms (CAINS) and the Brief Negative Symptom Assessment (BNSS) were developed to address conceptual and psychometric limitations with existing scales and fulfill recommendations from the NIMH-MATRICS consensus statement on negative symptoms,[36] which established five unique negative symptom factors: blunted affect, alogia, asociality, anhedonia, and avolition. The consensus group recommended the development of a new instrument that accurately measures the five agreed-upon factors. Additionally, the consensus statement recommended that the new tool also be able to distinguish between appetitive and consummatory anhedonia, or the frequency and quality of patient's social interactions vs the desire for such interactions. CAINS and BNSS have been used in clinical trials of negative symptoms, and both have demonstrated good psychometric properties. Both CAINS and BNSS have been translated into several languages, although validation studies for use in different countries and languages are largely pending for both scales.

CAINS

CAINS evaluates negative symptoms across 13 items divided into two main domains: Motivation and Pleasure (MAP) and Expression (EXP). The MAP domain assesses the individual's ability to experience pleasure and the motivation to engage in activities. It includes the assessment of anhedonia, asociality, and avolition. The EXP domain evaluates the individual's emotional expression and communication, and includes affective flattening and alogia. Ratings for each item are made on five-point scale of 0 (no impairment) to 4 (severe impairment).[47]

BNSS

BNSS also consists of 13 items and assesses symptoms across 6 domains: the five factors included in the CAINS as described above, plus one additional domain: lack of normal distress. BNSS uses a short semi-structed interview that can be administered in approximately 20 minutes. Ratings are based on information collected as well as clinician observations during the interview. Ratings for each item are made on a seven-point scale of 0 (absent) to 6 (severe impairment).[48]

Comparison of CAINS and BNSS

Both scales have been validated in more than one study involving a large sample population.[48-50] Advantages of CAINS include the explicit differentiation between the two primary domains, MAP and EXP, which have been found to be both robust and stable, providing useful information regarding the specific areas of impairment and differentiating them as potential treatment targets.[47] Confirmatory factor analyses conducted by Li et al. (2022)[51] of combined CAINS and BNSS data support the two factors as the latent structure for negative symptoms over the five-factor structure. Both CAINS and BNSS provide robust assessment of the five negative symptom factors. However, comparison of the psychometric properties of both scales shows low correlation for the anhedonia domains of the two scales, suggesting that the scales may differ on how this factor is measured [52]: for example, BNSS measures both intensity and frequency of pleasurable activities, whereas CAINS focuses only on frequency of pleasurable activities. On the other hand, CAINS provides a more nuanced assessment of range and

frequency of pleasurable activities in the anhedonia domain than BNSS.

There are also some noteworthy differences in the assessment of the asociality factor, which is split into two items in BNSS (vs one item in CAINS) allowing for a separate assessment of internal experience and overt behavior. The dissociation between these two constructs has been shown to predict clinical outcome.[53] Additionally, BNSS covers pathological lack of normal distress, which is a symptom frequently exhibited by patients with primary and enduring negative symptoms (i.e., deficit syndrome).[54] BNSS has shown higher test–retest reliability, and has demonstrated strong association with cognitive and neuropsychological testing.[52,55]

Regarding administration, CAINS can be more resource-intensive, taking 30–60 minutes to complete, whereas BNSS is quicker, and can be administered within 20 minutes. Both scales are supported by a manual with detailed instructions for administration and scoring; the CAINS manual provides sample vignettes, where the BNSS manual does not.

Finally, both scales have been criticized for neglecting to incorporate psychosocial and cognitive factors in schizophrenia.[37]

In summary, both CAINS and BNSS are valid tools for assessing negative symptoms in schizophrenia spectrum disorders in clinical trials, and the choice may depend on the study objectives and resources available. Similar to other instruments discussed, both scales require that clinicians are adequately trained on administration and rating procedures and have prior experience in the assessment of negative symptoms.

Clinical impression scales and assessments of functioning

CLINICAL GLOBAL IMPRESSION SCALES

The Clinical Global Impression (CGI) scale is a clinician-rated scale where raters are asked to provide a global impression of illness severity and overall functioning.[56] There are two CGI versions, frequently used together in the same study protocol: CGI-Severity (CGI-S) and CGI-Improvement (CGI-I), sometimes also referred to as CGI-Change (CGI-C). CGI-S rates the current level of psychopathology, whereas CGI- I/C rates the degree of improvement or change since the beginning of a specific treatment or intervention period. Typically, CGI is rated on 1–7 scale, although variations exist, and scales can be adapted to meet specific study objectives. CGI requires no interviews, and is rated based on the clinician's impression of the subject and considering all available information.

Advantages and Strengths

CGI is quick and simple to rate, relying on the clinician's judgment of severity or improvement. These scales can be customized to meet specific study needs. It also can be made culturally sensitive by integrating cultural beliefs and factors that may influence the individual's symptoms and functioning.

Disadvantages and Critique

Because it relies heavily on the clinician's impression, CGI requires significant clinician experience with the full spectrum of disease severity; it can also be subjective and prone to bias. CGI-I/C relies exclusively on the ability of the clinician to be able to remember the subject's baseline level of severity, which can be challenging for clinicians who assess multiple subjects in the course of a week in studies where the treatment assessment period is long.

Clinician Perspective

CGI is very quick and intuitive as rating scale, allowing clinicians to provide their holistic view of a subject's illness severity and functioning and progress in treatment.

Clinical Global Impression-Schizophrenia

Clinical Global Impression-Schizophrenia (CGI-SCH) was developed to separately assess clinical severity and treatment response in a subset of symptoms associated with schizophrenia: positive, negative, depressive, and cognitive.[57] This scale has not been widely used in clinical trials of psychosis, which generally favor use of the standard CGI-S/I, but may prove suitable for future use as it potentially distinguishes between domain-specific severity and targeted treatment response.

Global Assessment of Functioning Scale

The Global Assessment of Functioning (GAF) scale measures a person's overall level of psychological, social, and occupational functioning through a numerical rating system, ranging from 0–100; the higher the score, the better the individual's functioning.[58] Originally developed in 1962, it has been revised over time. While not specifically developed to assess psychosis, it can be a valuable tool in helping clinicians assess the impact of psychosis on the patient's life.

Advantages

GAF is a simple and quick assessment, and is known to many clinicians. Like CGI, GAF can be quickly rated without need of additional interviews.

Disadvantages and Critique

The GAF scale has been criticized for being overly subjective, for lacking comprehensive guidelines for ratings, and for its questionable validity and reliability.[59] For this reason, it is not frequently used currently in clinical trials.

Clinician Perspective

Selecting the exact measurement within a scale of 0–100 can be challenging and lack precision, particularly for less-experienced clinicians.

PERSONAL AND SOCIAL PERFORMANCE SCALE

The Personal and Social Performance (PSP) scale is a valid a reliable measure to assess social and occupational functioning in individuals in the acute or stable stage of schizophrenia.[60] The PSP score is widely used as a secondary endpoint to assess change in a subject's functioning in clinical trials in psychosis.

PSP score is based on the patient's performance in four domains: socially useful activities, personal and social relationships, self-care, and disturbing and aggressive behavior. PSP produces a total score of 0–100, divided in 10-point increments, which aims to provide a holistic view of the patient's global functioning.

Advantages and Strengths

A unique advantage is that PSP provides an objective measure of functioning across several domains. PSP has demonstrated sensitivity to change in functioning in studies in psychosis,[61-63] and has been shown to correlate well with PANSS.[64]

Disadvantages

PSP's complex scoring system can be difficult for clinicians, and interrater agreement has been reported to be low.[65]

Clinician Perspective

PSP requires that clinicians make subtle decisions regarding level of impairment in different areas of functioning. The final rating on the scale from 0–100 can be difficult to determine and, like GAF, can lack precision as clinicians try to choose precise ratings within the 10-point intervals.

QUALITY OF LIFE SCALE

Developed by Heinrichs et al. in 1984, the Quality-of-Life Scale (QLS) assesses quality of life in individuals with schizophrenia.[66] QLS consists of a semi-structured 21-item interview covering various domains associated with quality of life, such as physical health, psychological well-being, social relationships, and environmental factors. It is the most widely used scale to measure quality of life in psychosis research.

Advantages and Strengths

Unlike other widely used quality of life scales, QLS was designed specifically to measure quality of life in schizophrenia. It has been validated for use in multiple languages.

Disadvantages and Critique

The interview questions are outdated (e.g., asking about frequency of buying stamps or writing a letter), likely limiting its current and future validity. Additionally, the scale relies purely on the clinician's assessment and obviates the patient's subjective experience.

Clinician Perspective

QLS is somewhat time-consuming to administer, averaging close to 30 minutes, particularly after other time-consuming outcome assessments (e.g., PANSS). Subjects with psychosis often find it difficult to provide precise responses to the interview questions, increasing rater effort. As previously mentioned, raters can find it difficult to navigate the outdated interview prompts that do not apply to most subjects' current day to day experiences.

Assessment of depression in schizophrenia

CALGARY DEPRESSION SCALE FOR SCHIZOPHRENIA

The Calgary Depression Scale for Schizophrenia (CDSS) is a brief measure of depression for use in schizophrenia[67]; it includes nine items related to mood, feelings of guilt, hopelessness, and other symptoms commonly associated with depression. Items are scored on a 0–3 scale, from absent (0) to severe (3).

Advantages and Strengths

CDSS was developed specifically to assess depressive symptoms in individuals with schizophrenia; it has been found to be superior to the Hamilton Depression Rating Scale and Montgomery–Åsberg Scale for the use in this population, differentiating depressive symptoms from positive, negative, and extrapyramidal symptoms.[67-69] CDSS can be a useful tool when a more detailed assessment of depressive symptoms is required, as depression frequently co-occurs with psychotic disorders and can be difficult to separate from other symptoms of psychosis.

Disadvantages and Critique

As many other scales described here, CDSS must be administered by an experienced and trained clinician.

Clinician Perspective

CDSS is easy to administer and score. The limited four anchor-point approach enables the clinician to make a quick assessment for each item. Items such as guilty ideas of reference and pathological guilt do not provide explicit guidance on how to score when a subject is delusional. Training is encouraged to guide clinicians on rating these items.

Assessing cognitive ability in psychotic disorders

Although psychotic symptoms are often visibly disruptive and treated as a primary challenge for patients with schizophrenia, cognitive disturbances have been a more recent treatment target in schizophrenia research. Cognitive deficits are detectible early in onset, are persistent over the course of illness, and have some level of impact on almost all patients with schizophrenia.[70,71] Leading academic and industry experts have recommended that cognitive ability be assessed in subjects that are clinically stable in order to prevent potential confounding factors (like interfering psychotic symptoms).[72] Cognitive ability has been measured in too numerous of ways to be covered in this review; however we will discuss the "gold standard" for measuring cognitive ability within clinical research, as well as alternate performance- and interview-based methods for measuring cognitive dysfunction in schizophrenia spectrum disorders. Each of the assessments identified below have been translated and culturally adapted for use in multinational clinical trials.

THE MATRICS CONSENSUS COGNITIVE BATTERY

The MATRICS Consensus Cognitive Battery (MCCB) is the result of an initiative between industry leaders, academic experts, and government regulators to identify not only the challenges to reliably measuring cognition in schizophrenia, but also to detail a path for designing and demonstrating efficacy of cognitive-enhancing treatment options in schizophrenia.[73,74] The MCCB consists of ten subtests that were identified through the RAND/UCLA Appropriateness Method and selected based upon ideal psychometric properties, and, for some subtests, the availability of alternate forms. The ten subscales measure ability across seven cognitive domains (speed of processing, verbal learning, working memory, reasoning and problem solving, visual learning, social cognition, and attention and vigilance), and normative data exists to compute a composite T-score representing overall level of cognitive impairment in relation to a normative sample, controlling for demographic factors like age and sex. The composite score shows the strongest sensitivity to cognitive impairment, and is highly related to vocational outcomes.[75]

Advantages of the MCCB

As the MCCB is considered the "gold standard" for cognitive assessment in schizophrenia, the primary advantage of this battery is the ample support in the literature of its strong psychometric properties, demonstrating high test–retest reliability and small practice effects across clinical research settings.[76,77]

Disadvantages

Although the MCCB is shorter than a typical neuropsychological evaluation, with an estimated completion time of 60–90 minutes with breaks, and the subtests were chosen with an eye toward ease of administration, one common complaint from clinical staff is the duration of the MCCB assessment. The battery consists of ten individual subtests with unique administration rules and discontinuation criteria, which can prove challenging for inexperienced raters. Proper administration relies heavily on the rater to administer both accurately and efficiently to ensure the patient is giving their best effort on each assessment.

Another complaint is that the social cognitive task included in the MCCB (the Mayer-Salovey-Caruso Emotional Intelligence Test [MSCEIT]) may be limited in assessing the broad social cognitive deficits seen in schizophrenia,[78] and that its standard administration may be limited in cross-cultural adaptation (see Hellemann et al. [2017][79] for potential scoring adaptations). Thus, researchers may choose to use only the "Neurocognitive Composite" T-score, calculated using the scores from the nine non-social cognitive subtests, as a primary endpoint. Specific social cognitive batteries have been proposed,[80,81] though more work is needed in this area to develop valid cross-cultural scales.[82]

Clinician Perspective

The MCCB provides an extensive assessment of cognitive ability across multiple domains that have been identified as key deficit areas in schizophrenia. However, this is a primarily paper assessment with some subtests requiring props. This can be burdensome on raters as they must remember specific and varied administration criteria, and for raters with poor administration, there can be a significant impact on quality data collection.

BRIEF ASSESSMENT OF COGNITION IN SCHIZOPHRENIA

Developed in response to complaints about lengthy neurocognitive batteries, the Brief Assessment of

Cognition in Schizophrenia (BACS) can efficiently assess multiple deficits across multiple cognitive domains in schizophrenia in around 35 minutes[83] Although initially developed as a pen-and-paper neurocognitive assessment, BACS has been adapted into a digital, tablet-based, app-driven assessment that results in psychometric properties consistent with the paper form. The BACS battery consists of six subtests assessing performance across six key cognitive domains (verbal memory and learning, working memory, motor function, verbal fluency, speed of processing, and executive function). Similar to the MCCB, BACS subtests have alternate forms to allow for repeated administration, and the composite score demonstrates high test–retest reliability. BACS also has an available normative dataset to allow computation of a composite T-score representing overall level of cognitive impairment in relation to a normative sample, controlling for demographic factors like age and sex.

Advantages of BACS

As noted above, one key advantage of BACS assessment is the duration of the assessment: an estimated 30 minutes, roughly one-third of that of MCCB, the current "gold standard" in schizophrenia. Further to this, digital administration via the BACS app provides significant advantages in multi-site clinical trials. The BACS app utilizes standardized verbal instructions to reduce inter-rater variability and encourage uniform administration, and includes automatic scoring to reduce simple scoring errors. However, a local rater is still required to be present to ensure the subject is fully engaged and putting forth maximal effort.

Disadvantages of BACS

One major concern is that BACS is not fully analogous to the MCCB, as it does not contain subtests that assess the social cognitive, visual learning, or attention and vigilance cognitive domains. As the NIMH-MATRICS initiative identified these as distinct, separable factors,[84] there may be concern that BACS does not fully assess the spectrum of potential cognitive deficits, and thus cognitive improvements, that can be seen in schizophrenia spectrum disorders. This concern can be overcome by including supplemental assessments into trial designs as needed.

Clinician Perspective

Digital administration via the BACS app provides a more succinct assessment of cognition across key domains for patients with schizophrenia, and reduces some of the administrative burdens for the rater. Administration is standardized through a digital delivery, though a trained rater is still required to administer the assessment.

Functionally meaningful co-primary endpoints

For any (potentially) cognition-enhancing intervention, a demonstration of improvement in specific cognitive performance measures is necessary. However, improvement is not sufficient to prove efficacy unless a meaningful improvement in day-to-day cognitive functioning is also demonstrated.[74] Thus, any discussion on cognitive assessment should also include a discussion on a functionally meaningful co-primary endpoint. These endpoints are typically assessed via interview, or according to a measure of functional capacity. Examples for each of these types of assessments are discussed below.

SCHIZOPHRENIA COGNITION RATING SCALE

The Schizophrenia Cognition Rating Scale (SCoRS) is an interview-based method for assessment of cognitive ability that focuses on the day-to-day experience of the patient over the two weeks prior to the interview.[85] The interview utilizes 20 items to assess all seven MATRICS cognitive domains, which are then clinician-rated from 1 (none) to 4 (severe). Importantly, the clinician-led interview is conducted directly with the patient, as well as with an informant to support the clinician's judgement of the most accurate rating of cognitive difficulty. The interview is estimated to take 20–25 minutes to complete.

Advantages

SCoRS has demonstrated sound psychometric properties both in academic research and clinical trials,[86] with high test–retest reliability and consistent correlations with cognitive performance and functional skills, as well as sensitivity to treatment.[85,86]

Disadvantages

The primary disadvantage of any interview-based assessment is that it is reliant on accurate reporting. As a function of schizophrenia spectrum disorders, the concern for poor metacognition and self-reporting of cognitive difficulties is addressed by the use of an informant during the interview process. Patient-report alone has limited relationship with performance[86,87]; however, less reliable informant quality can also impact sensitivity to treatment effect.[86] Thus, there is a strong reliance on the available informant to provide a quality report on the day-to-day difficulties experienced by the patient. Further to this, the semi-structured nature of SCoRS requires the clinician to known when to probe

further to determine accurate ratings for any given item.

Clinician Perspective

The SCoRS assessment can be easily implemented in many diverse settings, and captures the perceived impact of cognitive difficulties on the subject's activities of daily living.

VIRTUAL REALITY FUNCTIONAL CAPACITY ASSESSMENT TOOL

A substantial challenge to measuring functional improvements within the confines of a clinical trial is that many ideal outcomes are limited by opportunity (e.g., maintaining employment, financial independence). By using a proxy measure of functional capacity, a patient's ability to perform certain independent activities can be assessed without relying on the individual to have certain circumstances available. The Virtual Reality Functional Capacity Assessment Tool (VRFCAT) is a digital assessment of functional capacity, allowing the patient to demonstrate their ability to efficiently perform, for example, a grocery shopping task.[88] The assessment is tablet-delivered and guides the patient through a series of tasks within a storyboard game-like environment. The final endpoint, the total amount of time to complete 12 separate task objectives, can be interpreted via comparisons to a normative dataset as an age and sex-based T-score; administration usually takes 25–30 minutes to completion.

Advantages

The major advantage of VRFCAT vs other assessments of functional capacity is digital administration. Other assessments of functional capacity require the use of props or role-playing that may present difficulties in site-level administration.[89] VRFCAT is wholly contained within the tablet environment, and instructions are uniform across all sites and administrations. A rater is still required to be present for administration to ensure the subject is fully engaged and putting forth maximal effort. Alternate forms allow for repeated assessment, and the scale has demonstrated excellent psychometric properties (i.e., limited practice effects, good test–retest reliability, and sensitivity to group differences at similar levels to the MCCB).[88]Further to this, VRFCAT has demonstrated sensitivity to treatment response[90] and cognitive decline[91]

Disadvantages

A primary concern is that VRFCAT is somewhat limited in the functional task that it assesses. As a shopping task, it does not directly assess other potentially important functional domains such as communication and social skills or medication management that are included in other assessments of functional capacity.[92-94] However, content analyses have revealed that VRFCAT seems to assess key domains important for functional independence (i.e., efficiency in using transportation and handling money).[95] A further concern is that administration requires the availability of specified devices.

Clinician Perspective

VRFCAT is well tolerated and removes the rater administration burden associated with other measures of functional capacity. The assessment is fully automated, and walks the subject through the objectives with standardized instructions. However, it does require a trained rater to monitor, ensuring that the subject is engaged and putting forth their best effort, which can be dull for raters as it usually takes 25–30 minutes for the subject to complete the task.

Beyond clinical outcome assessments: the future of psychosis assessment

There have been numerous articles and reviews on the limitations of subjective symptom evaluation in research and treatment of psychotic disorders, particularly in areas such as negative symptoms[96] and symptom clusters, which involve data from multiple observers.[97] The time-consuming nature of clinical interviewing, not to mention the persistent challenges of reliability and training, have made it extremely difficult to translate many commonly used research tools to routine clinical practice.[98]

Alternatives to clinical interviews and self-report questionnaires have been proposed, including the use of computerized tools such as automated analysis of language and speech characteristics, the use of ambulatory "digital phenotyping" to assess activity levels, movement, location, and other variables via smartphones, and ecological momentary assessment which allow "in the moment" self-report of subjective states and functioning.[99,100] Digital assessments lend themselves well for adoption into these newer technologies, as concurrent audio recordings can be collected from clinical interviews, and metadata from virtual reality or digital cognitive tests can be analyzed with more sophisticated exploratory data analytics. These innovations are promising, and some may help to provide a bridge between our current state of the art and a future where quantifiable, objective phenomena are related to meaningful outcomes. Furthermore, these novel methods may clarify the interplay between various symptoms observed in schizophrenia spectrum disorders, identifying potentially common treatment targets or

identifying subgroups that are uniquely receptive to specific treatments.

It is important to remember, as we work towards a future state where mobile devices, virtual reality tests, and automated assessment of face and voice become accepted as clinically valid, that our existing rating scales, diagnostic instruments, and psychometric tools did not arise in a vacuum. They are products of their time, with input from culture, biases, and the values of the times in which they were developed. As the next generation of tools evolves, they too will ultimately have their strengths and drawbacks, biases, and challenges. It will be the responsibility of every researcher and clinician to partner with advocates and ensure that they are helping to advance the art and science of psychosis research and treatment.

Conflicts of interest statement

All authors are employees of WCG Clinical, Inc. Mark Opler receives royalties from Pearson, Inc.

References

1. Walton MK, Powers JH 3rd, Hobart J, et al. Clinical outcome assessments: conceptual foundation-report of the ISPOR clinical outcomes assessment - emerging good practices for Outcomes Research Task Force. *Value Health*. 2015;18(6):741-752. doi: 10.1016/j.jval.2015.08.006
2. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59 Suppl 20:22-57.
3. First MB, Williams JBW, Karg RS, Spitzer RL. Structured Clinical Interview for DSM-5—Research Version (SCID-5 for DSM-5, Research Version; SCID-5-RV). Arlington, VA, American Psychiatric Association, 2015.
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*. American Medical Association; 2022.
5. World Health Organization (WHO). *The ICD-10 classification of mental and behavioural disorders*. World Health Organization; 1993
6. Gjerden P. Criteria-based diagnostics generate excessive and incorrect diagnoses. *Tidsskr Nor Laegeforen*. 2015;135(16):1437. doi: 10.4045/tidsskr.15.0727
7. Shabani A, Masoumian S, Zamirinejad S, et al. Psychometric properties of Structured Clinical Interview for DSM-5 Disorders-Clinician Version (SCID-5-CV). *Brain Behav*. 2021;11(5):e01894. doi: 10.1002/brb3.1894
8. Osório FL, Loureiro SR, Hallak JEC, et al. Clinical validity and intrarater and test-retest reliability of the Structured Clinical Interview for DSM-5 - Clinician Version (SCID-5-CV). *Psychiatry Clin Neurosci*. 2019;73(12):754-760. doi: 10.1111/pcn.12931
9. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261-276. doi: 10.1093/schbul/13.2.261
10. Marder SR, Davis JM, Chouinard G. The effects of risperidone on the five dimensions of schizophrenia derived by factor analysis: combined results of the North American trials [published correction appears in J Clin Psychiatry 1998 Apr;59(4):200]. *J Clin Psychiatry*. 1997;58(12):538-546. doi: 10.4088/jcp.v58n1205
11. Hopkins SC, Ogirala A, Loebel A, Koblan KS. Transformed PANSS factors intended to reduce pseudospecificity among symptom domains and enhance understanding of symptom change in antipsychotic-treated patients with schizophrenia. *Schizophr Bull*. 2018;44(3):593-602. doi: 10.1093/schbul/sbx101
12. Opler LA, Kay SR, Lindenmayer JP, Fiszbein A. SCI- PANSS. Toronto: Multi-Health Systems Inc; 1992.
13. Khan A, Yavorsky C, Liechti S, et al. A Rasch model to test the cross-cultural validity in the positive and negative syndrome scale (PANSS) across six geo-cultural groups. *BMC Psychol*. 2013;1(1):5. doi: 10.1186/2050-7283-1-5
14. Kay SR, Opler LA, Lindenmayer JP. Reliability and validity of the positive and negative syndrome scale for schizophrenics. *Psychiatry Res*. 1988;23(1):99-110. doi: 10.1093/schbul/13.2.261
15. Santor DA, Ascher-Svanum H, Lindenmayer JP, Obenchain RL. Item response analysis of the Positive and Negative Syndrome Scale. *BMC Psychiatry*. 2007;7:66. doi: 10.1186/1471-244X-7-66
16. Lindström E, Wieselgren IM, von Knorring L. Interrater reliability of the Structured Clinical Interview for the Positive and Negative Syndrome Scale for schizophrenia. *Acta Psychiatr Scand*. 1994;89(3):192-195. doi: 10.1111/j.1600-0447.1994.tb08091.x
17. Bentsen H, Munkvold O, Notland TH, et al. The interrater reliability of the Positive and Negative Syndrome Scale (PANSS). *Int J Methods Psychiatr Res*. 1996;6(4):227-235.
18. Nielsen CM, Kølbæk P, Dines D, Opler M, Correll CU, Østergaard SD. Are informants required to obtain valid ratings on the Positive and Negative Syndrome Scale (PANSS)?. *Schizophrenia (Heidelb)*. 2023;9(1):54. doi: 10.1038/s41537-023-00378-5
19. Opler MGA, Yavorsky C, Daniel DG. Positive and Negative Syndrome Scale (PANSS) training: challenges, solutions, and future directions. *Innov Clin Neurosci*. 2017;14(11-12):77-81.
20. Østergaard SD, Lemming OM, Mors O, Correll CU, Bech P. PANSS-6: a brief rating scale for the measurement of severity in schizophrenia. *Acta Psychiatr Scand*. 2016;133(6):436-444.
21. Nielsen CM, Kølbæk P, Dines D, et al. Validation of ratings on the six-item Positive and Negative Syndrome Scale obtained via the Simplified Negative and Positive Symptoms Interview among outpatients with schizophrenia. *J Psychopharmacol*. 2022;36(11):1208-1217. doi: 10.1177/02698811221131992
22. Kølbæk P, Dines D, Holm T, et al. Clinical validation of ratings on the six-item Positive and Negative Syndrome Scale obtained via the Simplified Negative and Positive Symptoms Interview. *J Psychopharmacol*. 2021;35(9):

1081-1090. doi: 10.1177/026988112199689
23. Østergaard SD, Foldager L, Mors O, Bech P, Correll CU. The validity and sensitivity of PANSS-6 in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study. *Schizophr Bull*. 2018;44(2):453-462. doi: 10.1093/schbul/sbx076
24. Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychol Rep*. 1962;10:799-812. doi: 10.2466/pr0.1962.10.3.799
25. Overall JE, Gorham DR. Brief Psychiatric Rating Scale (BPRS): recent developments in ascertainment and scaling. *Psychopharmacol Bull*. 1988;24(1):97-99.
26. Ventura J, Lukoff D, Nuechterlein KH, Liberman RP, Green M, Shaner A. Appendix 1: Brief Psychiatric Rating Scale (BPRS) Expanded version (4.0) scales, anchor points and administration manual. *Int J Methods Psychiatr Res*. 1993;3:227-244.
27. Hofmann AB, Schmid HM, Jabat M, et al. Utility and validity of the Brief Psychiatric Rating Scale (BPRS) as a transdiagnostic scale. *Psychiatry Res*. 2022;314:114659. doi: 10.1016/j.psychres.2022.114659
28. Zanello A, Berthoud L, Ventura J, Merlo MC. The Brief Psychiatric Rating Scale (version 4.0) factorial structure and its sensitivity in the treatment of outpatients with unipolar depression. *Psychiatry Res*. 2013;210(2):626-633. doi: 10.1016/j.psychres.2013.07.001
29. Rhoades HM, Overall JE. The semistructured BPRS interview and rating guide. *Psychopharmacol Bull*. 1988;24(1):101-104.
30. Schooler NR. Precursors to the PANSS: the BPRS and its progenitors. *Innov Clin Neurosci*. 2017;14(11-12):10-11.
31. Andreasen NC. Scale for the Assessment of Positive Symptoms (SAPS). 1986. Accessed November 15, 2023. https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft48377-000
32. Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br J Psychiatry Suppl*. 1989;(7):49-58.
33. The Parkinson Study Group. Low-dose clozapine for the treatment of drug-induced psychosis in Parkinson's disease. *N Engl J Med*. 1999;340(10):757-763. doi: 10.1056/NEJM199903113401003
34. Isaacson SH, Coate B, Norton J, Stankovic S. Blinded SAPS-PD assessment after 10 weeks of pimavanserin treatment for Parkinson's disease psychosis. *J Parkinsons Dis*. 2020;10(4):1389-1396. doi: 10.3233/JPD-202047
35. Isaacson SH, Ballard CG, Kreitzman DL, et al. Efficacy results of pimavanserin from a multi-center, open-label extension study in Parkinson's disease psychosis patients. *Parkinsonism Relat Disord*. 2021;87:25-31. doi: 10.1016/j.parkreldis.2021.04.012
36. Kirkpatrick B, Fenton WS, Carpenter WT Jr, Marder SR. The NIMH-MATRICS consensus statement on negative symptoms. *Schizophr Bull*. 2006;32(2):214-219. doi: 10.1093/schbul/sbj053
37. Kumari S, Malik M, Florival C, Manalai P, Sonje S. An assessment of five (PANSS, SAPS, SANS, NSA-16, CGI-SCH) commonly used symptoms rating scales in schizophrenia and comparison to newer scales (CAINS, BNSS). *J Addict Res Ther*. 2017;8(3):324. doi: 10.4172/2155-6105.1000324
38. Alphs LD, Summerfelt A, Lann H, Muller RJ. The negative symptom assessment: a new instrument to assess negative symptoms of schizophrenia. *Psychopharmacol Bull*. 1989;25(2):159-163.
39. Axelrod BN, Goldman RS, Alphs LD. Validation of the 16-item Negative Symptom Assessment. *J Psychiatr Res*. 1993;27(3):253-258. doi: 10.1016/0022-3956(93)90036-2
40. Eckert SL, Diamond PM, Miller AL, Velligan DI, Funderburg LG, True JE. A comparison of instrument sensitivity to negative symptom change. *Psychiatry Res*. 1996;63(1):67-75. doi: 10.1016/0165-1781(96)02870-3
41. Rekhi G, Alphs L, Ang MS, Lee J. Clinical utility of the Negative Symptom Assessment-16 in individuals with schizophrenia. *Eur Neuropsychopharmacol*. 2019;29(12):1433-1441. doi: 10.1016/j.euroneuro.2019.10.009
42. Velligan DI, Lopez L, Castillo DA, Manaugh B, Milam AC, Miller AL. Interrater reliability of using brief standardized outcome measures in a community mental health setting. *Psychiatr Serv*. 2011;62(5):558-560. doi: 10.1176/ps.62.5.pss6205_0558
43. Daniel DG, Alphs L, Cazorla P, Bartko JJ, Panagides J. Training for assessment of negative symptoms of schizophrenia across languages and cultures: comparison of the NSA-16 with the PANSS Negative Subscale and Negative Symptom factor. *Clin Schizophr Relat Psychoses*. 2011;5(2):87-94. doi: 10.3371/CSRP.5.2.5
44. Blanchard JJ, Kring AM, Horan WP, Gur R. Toward the next generation of negative symptom assessments: the collaboration to advance negative symptom assessment in schizophrenia. *Schizophr Bull*. 2011;37(2):291-299. doi: 10.1093/schbul/sbq104
45. Alphs L, Morlock R, Coon C, van Willigenburg A, Panagides J. The 4-item Negative Symptom

Assessment (NSA-4) instrument: a simple tool for evaluating negative symptoms in schizophrenia following brief training. *Psychiatry (Edgmont)*. 2010;7(7):26-32.

46. Alphs L, Morlock R, Coon C, Cazorla P, Szegedi A, Panagides J. Validation of a 4-item Negative Symptom Assessment (NSA-4): a short, practical clinical tool for the assessment of negative symptoms in schizophrenia. *Int J Methods Psychiatr Res*. 2011;20(2):e31-e37. doi: 10.1002/mpr.339
47. Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. *Am J Psychiatry*. 2013;170(2):165-172. doi: 10.1176/appi.ajp.2012.12010109
48. Kirkpatrick B, Strauss GP, Nguyen L, et al. The Brief Negative Symptom Scale: psychometric properties. *Schizophr Bull*. 2011;37(2):300-305. doi: 10.1093/schbul/sbq059
49. Carpenter WT, Blanchard JJ, Kirkpatrick B. New standards for negative symptom assessment. *Schizophr Bull*. 2016;42(1):1-3. doi: 10.1093/schbul/sbv160
50. Strauss GP, Keller WR, Buchanan RW, et al. Next-generation negative symptom assessment for clinical trials: validation of the Brief Negative Symptom Scale. *Schizophr Res*. 2012;142(1-3):88-92. doi: 10.1016/j.schres.2012.10.012
51. Li SB, Liu C, Zhang JB, et al. Revisiting the latent structure of negative symptoms in schizophrenia: evidence from two second-generation clinical assessments [published correction appears in Schizophr Res. 2023 Feb;252:206-207]. *Schizophr Res*. 2022;248:131-139. doi: 10.1016/j.schres.2022.08.016
52. Strauss GP, Gold JM. A psychometric comparison of the Clinical Assessment Interview for Negative Symptoms and the Brief Negative Symptom Scale. *Schizophr Bull*. 2016;42(6):1384-1394. doi: 10.1093/schbul/sbw046
53. Strauss GP, Horan WP, Kirkpatrick B, et al. Deconstructing negative symptoms of schizophrenia: avolition-apathy and diminished expression clusters predict clinical presentation and functional outcome. *J Psychiatr Res*. 2013;47(6):783-790. doi: 10.1016/j.jpsychires.2013.01.015
54. Kirkpatrick B, Buchanan RW, Ross DE, Carpenter WT Jr. A separate disease within the syndrome of schizophrenia. *Arch Gen Psychiatry*. 2001;58(2):165-171. doi: 10.1001/archpsyc.58.2.165
55. Strauss GP, Keller WR, Buchanan RW, et al. Next-generation negative symptom assessment for clinical trials: validation of the Brief Negative Symptom Scale. *Schizophr Res*. 2012;142(1-3):88-92. doi: 10.1016/j.schres.2012.10.012
56. Guy W. Clinical Global Impressions Scale (CGI). 1976. Accessed November 15, 2023, https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft48216-000
57. Haro JM, Kamath SA, Ochoa S, et al. The Clinical Global Impression-Schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia. *Acta Psychiatr Scand Suppl*. 2003;(416):16-23. doi: 10.1034/j.1600-0447.107.s416.5.x
58. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). American Psychiatric Association; 2000.
59. Aas IH. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. *Ann Gen Psychiatry*. 2010;9:20. doi: 10.1186/1744-859X-9-20
60. Morosini PL, Magliano L, Brambilla L, Ugolini S, Pioli R. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social functioning. *Acta Psychiatr Scand*. 2000;101(4):323-329.
61. Kane J, Canas F, Kramer M, et al. Treatment of schizophrenia with paliperidone extended-release tablets: a 6-week placebo-controlled trial. *Schizophr Res*. 2007;90(1-3):147-161. doi: 10.1016/j.schres.2006.09.012
62. Patrick DL, Burns T, Morosini P, et al. Reliability, validity and ability to detect change of the clinician-rated Personal and Social Performance scale in patients with acute symptoms of schizophrenia. *Curr Med Res Opin*. 2009;25(2):325-338. doi: 10.1185/03007990802611919
63. Nasrallah H, Morosini P, Gagnon DD. Reliability, validity and ability to detect change of the Personal and Social Performance scale in patients with stable schizophrenia. *Psychiatry Res*. 2008;161(2):213-224. doi: 10.1016/j.psychres.2007.11.012
64. Jelastopulu E, Giourou E, Merekoulias G, Mestousi A, Moratis E, Alexopoulos EC. Correlation between the Personal and Social Performance scale (PSP) and the Positive and Negative Syndrome Scale (PANSS) in a Greek sample of patients with schizophrenia. *BMC Psychiatry*. 2014;14:197. doi: 10.1186/1471-244X-14-197

65. White S, Dominise C, Naik D, Killaspy H. The reliability of the Personal and Social Performance scale - informing its training and use. *Psychiatry Res*. 2016;243:312-317. doi: 10.1016/j.psychres.2016.06.047
66. Heinrichs DW, Hanlon TE, Carpenter WT Jr. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. *Schizophr Bull*. 1984;10(3):388-398. doi:10.1093/schbul/10.3.388
67. Addington D, Addington J, Maticka-Tyndale E, Joyce J. Reliability and validity of a depression rating scale for schizophrenics. *Schizophr Res*. 1992;6(3):201-208. doi: 10.1016/0920-9964(92)90003-n
68. Addington D, Addington J, Maticka-Tyndale E. Assessing depression in schizophrenia: the Calgary Depression Scale. *Br J Psychiatry Suppl*. 1993;(22):39-44.
69. Kim SW, Kim SJ, Yoon BH, et al. Diagnostic validity of assessment scales for depression in patients with schizophrenia. *Psychiatry Res*. 2006;144(1):57-63. doi: 10.1016/j.psychres.2005.10.002
70. Harvey PD, Bosia M, Cavallaro R, et al. Cognitive dysfunction in schizophrenia: An expert group paper on the current state of the art. *Schizophr Res Cogn*. 2022;29:100249. doi: 10.1016/j.scog.2022.100249
71. Keefe RS, Eesley CE, Poe MP. Defining a cognitive function decrement in schizophrenia. *Biol Psychiatry*. 2005;57(6):688-691. doi: 10.1016/j.biopsych.2005.01.003
72. Marder SR. The NIMH-MATRICS project for developing cognition-enhancing agents for schizophrenia. *Dialogues Clin Neurosci*. 2006;8(1):109-113. doi: 10.31887/DCNS.2006.8.1/smarder
73. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry*. 2008;165(2):203-213. doi: 10.1176/appi.ajp.2007.07010042
74. Buchanan RW, Davis M, Goff D, et al. A summary of the FDA-NIMH-MATRICS workshop on clinical trial design for neurocognitive drugs for schizophrenia. *Schizophr Bull*. 2005;31(1):5-19. doi: 10.1093/schbul/sbi020
75. August SM, Kiwanuka JN, McMahon RP, Gold JM. The MATRICS Consensus Cognitive Battery (MCCB): clinical and cognitive correlates. *Schizophr Res*. 2012;134(1):76-82. doi: 10.1016/j.schres.2011.10.015
76. Keefe RS, Fox KH, Harvey PD, Cucchiaro J, Siu C, Loebel A. Characteristics of the MATRICS Consensus Cognitive Battery in a 29-site antipsychotic schizophrenia clinical trial. *Schizophr Res*. 2011;125(2-3):161-168. doi: 10.1016/j.schres.2010.09.015
77. Georgiades A, Davis VG, Atkins AS, et al. Psychometric characteristics of the MATRICS Consensus Cognitive Battery in a large pooled cohort of stable schizophrenia patients. *Schizophr Res*. 2017;190:172-179. doi: 10.1016/j.schres.2017.03.040
78. Eack SM, Greeno CG, Pogue-Geile MF, Newhill CE, Hogarty GE, Keshavan MS. Assessing social-cognitive deficits in schizophrenia with the Mayer-Salovey-Caruso Emotional Intelligence Test. *Schizophr Bull*. 2010;36(2):370-380. doi: 10.1093/schbul/sbn091
79. Hellemann GS, Green MF, Kern RS, Sitarenios G, Nuechterlein KH. Developing an international scoring system for a consensus-based social cognition measure: MSCEIT-managing emotions. *Psychol Med*. 2017;47(14):2494-2501. doi:10.1017/S0033291717001052
80. Pinkham AE, Harvey PD, Penn DL. Social cognition psychometric evaluation: results of the final validation study. *Schizophr Bull*. 2018;44(4):737-748. doi: 10.1093/schbul/sbx117
81. Brunet-Gouet E, Decaix-Tisserand C, Urbach M, et al. Outcome prediction with a social cognitive battery: a multicenter longitudinal study. *NPJ Schizophr*. 2021;7(1):30. doi: 10.1038/s41537-021-00160-5
82. Hajdúk M, Achim AM, Brunet-Gouet E, Mehta UM, Pinkham AE. How to move forward in social cognition research? Put it into an international perspective. *Schizophr Res*. 2020;215:463-464. doi: 10.1016/j.schres.2019.10.001
83. Keefe RS, Poe M, Walker TM, Harvey PD. The relationship of the Brief Assessment of Cognition in Schizophrenia (BACS) to functional capacity and real-world functional outcome. *J Clin Exp Neuropsychol*. 2006;28(2):260-269. doi: 10.1080/13803390500360539
84. Green MF, Nuechterlein KH, Gold JM, et al. Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICS conference to select cognitive domains and test criteria. *Biol Psychiatry*. 2004;56(5):301-307. doi: 10.1016/j.biopsych.2004.06.023
85. Keefe RS, Poe M, Walker TM, Kang JW, Harvey PD. The Schizophrenia Cognition Rating Scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *Am J Psychiatry*. 2006;163(3):426-432. doi: 10.1176/appi.ajp.163.3.42

86. Keefe RS, Davis VG, Spagnola NB, et al. Reliability, validity and treatment sensitivity of the Schizophrenia Cognition Rating Scale. *Eur Neuropsychopharmacol*. 2015;25(2):176-184. doi: 10.1016/j.euroneuro.2014.06.009
87. Balzan RP, Neaves A, Denson LA, Liu D, Galletly C. Cognitive deficit awareness in schizophrenia: absent, intact, or somewhere in-between?. *Cogn Neuropsychiatry*. 2014;19(6):471-484. doi: 10.1080/13546805.2014.909311
88. Keefe RSE, Davis VG, Atkins AS, et al. Validation of a computerized test of functional capacity. *Schizophr Res*. 2016;175(1-3):90-96. doi: 10.1016/j.schres.2016.03.038
89. Green MF, Schooler NR, Kern RS, et al. Evaluation of functionally meaningful measures for clinical trials of cognition enhancement in schizophrenia. *Am J Psychiatry*. 2011;168(4):400-407. doi: 10.1176/appi.ajp.2010.10030414
90. Nahum M, Lee H, Fisher M, et al. Online social cognition training in schizophrenia: a double-blind, randomized, controlled multi-site clinical trial. *Schizophr Bull*. 2021;47(1):108-117. doi: 10.1093/schbul/sbaa085
91. Atkins AS, Khan A, Ulshen D, et al. Assessment of instrumental activities of daily living in older adults with subjective cognitive decline using the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). *J Prev Alzheimers Dis*. 2018;5(4):216-234. doi: 10.14283/jpad.2018.28
92. Patterson TL, Goldman S, McKibbin CL, Hughs T, Jeste DV. UCSD Performance-Based Skills Assessment: development of a new measure of everyday functioning for severely mentally ill adults. *Schizophr Bull*. 2001;27(2):235-245. doi: 10.1093/oxfordjournals.schbul.a006870
93. Patterson TL, Moscona S, McKibbin CL, Davidson K, Jeste DV. Social skills performance assessment among older patients with schizophrenia. *Schizophr Res*. 2001;48(2-3):351-360. doi: 10.1016/s0920-9964(00)00109-2
94. Patterson TL, Lacro J, McKibbin CL, Moscona S, Hughs T, Jeste DV. Medication management ability assessment: results from a performance-based measure in older outpatients with schizophrenia. *J Clin Psychopharmacol*. 2002;22(1):11-19. doi: 10.1097/00004714-200202000-00003
95. Horan WP, Depp CA, Hurst S, et al. Qualitative analysis of the content validity of the Virtual Reality Functional Capacity Assessment Tool (VRFCAT) in schizophrenia: a multi-stakeholder perspective. *Schizophr Bull*. 2023;, sgad012. doi: 10.1093/schizbullopen/sgad012
96. Gupta M, Holshausen K, Gou L, Bowie C. Measuring negative symptom change in schizophrenia: considering alternatives to self-report. *Expert Rev Neurother*. 2014;14(8):911-922. doi: 10.1586/14737175.2014.935341
97. Ho BC, Flaum M, Hubbard W, Arndt S, Andreasen NC. Validity of symptom assessment in psychotic disorders: information variance across different sources of history. *Schizophr Res*. 2004;68(2-3):299-307. doi: 10.1016/j.schres.2003.07.006
98. Kobak KA, Engelhardt N, Williams JB, Lipsitz JD. Rater training in multicenter clinical trials: issues and recommendations. *J Clin Psychopharmacol*. 2004;24(2):113-117. doi: 10.1097/01.jcp.0000116651.91923.54
99. Cohen AS, Alpert M, Nienow TM, Dinzeo TJ, Docherty NM. Computerized measurement of negative symptoms in schizophrenia. *J Psychiatr Res*. 2008;42(10):827-836. doi: 10.1016/j.jpsychires.2007.08.008
100. Strauss GP, Zamani-Esfahlani F, Raugh IM, Luther L, Sayama H. Network analysis of discrete emotional states measured via ecological momentary assessment in schizophrenia [published online ahead of print, 2023 Jun 6]. *Eur Arch Psychiatry Clin Neurosci*. 2023;10.1007/s00406-023-01623-9. doi: 10.1007/s00406-023-01623-9