

Proxy measures for the assessment of psychotic and affective symptoms in studies using electronic health records

Álvaro López-Díaz, Fernanda Jazmín Palermo-Zeballos, Luis Gutierrez-Rojas, Luis Alameda, Francisco Gotor-Sánchez-Luengo, Nathalia Garrido-Torres, Johann Métrailler, Livia Alerci, Vincent Bonnard, Pablo Cano-Domínguez, Elma Avanesi-Molina, Miguel Soto-Ontoso, Rocio Torrecilla-Olavarrieta, Leticia Irene Muñoz-Manchado, Pedro Torres-Hernández, Fermín González-Higueras, Juan Luis Prados-Ojeda, Mario Herrera-Cortés, José Miguel Meca-García, Rafael Manuel Gordillo-Urbano, Cristina Sánchez-Robles, Tomás Delgado-Durán, María Felipa Soriano-Peña, Philippe Golay, Philippe Conus, Benedicto Crespo-Facorro and Miguel Ruiz-Veguilla

Background

There is a lack of standardised psychometric data in electronic health record (EHR)-based research. Proxy measures of symptom severity based on patients' clinical records may be useful surrogates in mental health EHR research.

Aims

This study aimed to validate proxy tools for the short versions of the Positive and Negative Syndrome Scale (PANSS-6), Young Mania Rating Scale (YMRS-6) and Montgomery-Åsberg Depression Rating Scale (MADRS-6).

Method

A cross-sectional, multicentre study was conducted in a sample of 116 patients with first-episode psychosis from 12 public hospitals in Spain. Concordance between PANSS-6, YMRS-6 and MADRS-6 scores and their respective proxies was evaluated based on information from EHR clinical notes, using a variety of statistical procedures, including multivariate tests to adjust for potential confounders. Bootstrapping techniques were used for internal validation, and an independent cohort from the Treatment and Early Intervention in Psychosis Program (TIPP-Lausanne, Switzerland) for external validation.

Results

The proxy versions correlated strongly with their respective standardised scales (partial correlations ranged from 0.75

to 0.84) and had good accuracy and discriminatory power in distinguishing between patients in and not in remission (percentage of patients correctly classified ranged from 83.9 to 91.4% and bootstrapped optimism-corrected area under the receiver operating characteristic curve ranged from 0.76 to 0.89), with high interrater reliability (intraclass correlation coefficient of 0.81). The findings remained robust in the external validation data-set.

Conclusions

The proxy instruments proposed for assessing psychotic and affective symptoms by reviewing EHR provide a feasible and reliable alternative to traditional structured psychometric procedures, and a promising methodology for real-world practice settings.

Keywords

Schizophrenia; psychosis; affective disorders; electronic health records; clinical assessment tool.

Copyright and usage

© The Author(s), 2024. Published by Cambridge University Press on behalf of the Royal College of Psychiatrists. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Periodic monitoring of psychotic and affective symptoms in the early stages of psychosis is essential for improving mental health outcomes in individuals with psychotic disorders.^{1,2} Regular assessment of those symptoms is also crucial for assessing the patient's clinical course over time, examining the efficacy of therapeutic interventions and identifying early signs of relapse.³ There are numerous assessment instruments for the severity of psychotic and affective symptoms, including self-report, structured and semi-structured clinician-rated scales.⁴ Among the clinician-rated scales for measuring psychotic symptoms in patients with schizophrenia and other psychotic disorders, the Positive and Negative Syndrome Scale (PANSS), Brief Psychiatric Rating Scale (BPRS) and Clinical Global Impression-Schizophrenia scale (CGI-SCH) are the most widely applied.^{5–7} The Young Mania Rating Scale (YMRS), Montgomery-Åsberg Depression Rating Scale (MADRS) and Hamilton Rating Scale for Depression are among the most widely used instruments for assessing the severity of manic and depressive symptoms in patients with bipolar disorder and depressive disorders.^{8–10} Outside of the research framework, and with the exception of the CGI-SCH (which is a simple, brief instrument for

evaluating severity and treatment response, but does not capture the full spectrum of psychotic features), all of these scales require significant administration and assessment time, as well as technical training for their reliable application. This usually makes them difficult to use in time-constrained settings such as clinical practice in community mental health services,¹¹ and is the main reason why these psychometric instruments are seldom used for monitoring symptom severity and treatment response in large-scale, real-world research, such as electronic health record (EHR)-based studies. Furthermore, there is growing concern about selection bias in psychosis research, as findings are usually based on a subset of patients who are willing to participate in research studies.¹²

A proxy measure is a tool used when it is not possible to directly assess a construct or phenomenon of interest.¹³ Proxy measures rely on the assumption that certain variables may be a good indicator of others that cannot be or are difficult to measure directly by standardised methods.¹³ Proxy tools have great potential in health science research, since they can provide an effective and efficient alternative for the measurement of outcomes or constructs that cannot be easily collected in daily practice. Several types of proxy indicators are used

in health sciences. One is psychometric assessment by close relatives or key informants (third-party assessments) when the patient cannot be directly assessed by the clinician, or when the patient is unable to report their own symptoms. Parent proxy-report versions of the MADRS and the YMRS are examples.^{14,15} Another type of proxy measure uses information from a surrogate scale replacing a more complex or time-consuming assessment tool. Some examples can be seen in the use of the National Adult Reading Test as a proxy indicator of IQ, or the 'Proxy for the Deficit Syndrome' (PDS), a case identification method that uses a set of the PANSS or the BPRS items to determine deficit schizophrenia status.^{16,17} Finally, in the third type of proxy measures, the construct of interest is assessed indirectly through demographic and socioeconomic variables. Some examples are residential postal codes as a measure of socioeconomic status; total healthcare costs as an indirect estimate of the intensity of health service use; or the sum of being unmarried, unemployed and having lower education as an indicator of poor premorbid psychosocial adjustment.^{18–20} The latter are particularly useful in reviewing medical records or in EHR-based studies. However, despite their potential, the number of validated proxies available is still scant, and their use in mental health research is not yet widespread.

Aims

The aim of this study was to develop and validate proxy measures of psychotic and affective symptoms based on the short versions of the PANSS, YMRS and MADRS scales, using unstructured clinical notes from EHRs. Such proxy measures would be very useful for assessing the severity of clinical features in retrospective chart reviews and in observational studies of real-world practice in patients with psychosis. They could, therefore, facilitate the implementation of large-scale studies based on representative samples of people with a psychotic disorder.

Method

Study population and inclusion/exclusion criteria

A multicentre, case-register study was conducted to examine the correlations between the short versions of the PANSS, YMRS and MADRS scales and their proxy versions. The sample was recruited from patients aged 18–55 years attending the mental health services (emergency, in-patient and out-patient settings) of 12 public hospitals in Andalusia (southern Spain) for treatment of first-episode psychosis (FEP). In addition, an independent cohort of patients from the Treatment and Early Intervention in Psychosis Program (TIPP-Lausanne; Switzerland) was used for external validation of the proxy instruments.²¹ The sum of all of these hospitals comprised an epidemiological catchment area of 5.5 million (approximately 3 million aged 18–55 years). Psychosis was confirmed with the Mini-International Neuropsychiatric Interview in the primary cohort and the Comprehensive Assessment of At Risk Mental States in the TIPP-Lausanne.^{22,23} All patients met the ICD-10 diagnostic criteria for psychotic disorders (F1x.4, F1x.5, F1x.7, F20–29, F30.2, F31.2, F31.5, F31.6, F32.3, F33.3 and F53.1 codes).²⁴ Exclusion criteria were organic psychoses, past diagnosis of psychotic disorder, severe or unstable medical condition, history of traumatic brain injury, cognitive impairment or major neurological disease, or any level of intellectual disability or autism spectrum disorder. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement guidelines were followed (see checklist in Supplementary Table 1 available at <https://doi.org/10.1192/bjo.2023.623> in Supplementary Appendix 1).²⁵ The authors assert that all procedures contributing to this work comply with

the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Patient consent was obtained in accordance with the requirements of the Andalusian Biomedical Research Ethics Committee (approval number 2184-N-21; informed consent was not required because of the use of de-identified data) and the Human Research Ethics Committee of the Canton Vaud (protocol number 2020-00272; verbal consent was witnessed and formally recorded). All procedures involving human participants in this research were approved by both ethics committees.

Selected demographic, clinical and psychometric variables

Information on patient sociodemographic, clinical and psychometric characteristics was acquired at study enrolment by a clinician-administered questionnaire designed by the authors. Sociodemographic data collected included age, gender, ethnicity (European White or other), marital status (married/partnership or unmarried), living situation (alone or with others), education (higher education and secondary/lower education) and occupation (employed, unemployed or student). Clinical characteristics collected included age at onset of psychosis, duration of untreated psychosis (following the methodology outlined in the Nottingham Onset Schedule²⁶), first-degree family history of psychosis, premorbid psychosocial adjustment (assessed with a validated proxy measure to discriminate patients with from those without poor premorbid adjustment²⁰), psychiatric history, history of harmful or hazardous substance use (alcohol, cannabis or other illicit drugs), type of psychotic disorder (according to ICD-10 coding), presence of suicidal symptoms (including suicidal ideation, plans or attempts according to the Paykel Suicide Scale methodology^{27–29}) and patient status (in-patient *v.* out-patient) at the time of enrolment. Psychometric measures were assessed at time of enrolment, using the shortened versions of the PANSS, YMRS and MADRS and their respective proxies.^{30–32} The time of the psychometric assessment (acute or stabilisation phase) and the experience of the clinician assessing the proxies (less than or over 5 years) were also recorded.

Abbreviated versions of PANSS, YMRS and MDRS scales

The six-item version of the PANSS (PANSS-6) is a brief rating scale focusing on the severity of key schizophrenia symptoms. It consists of three items each for the positive and negative subscales (P1 delusions, P2 conceptual disorganisation, P3 hallucinations, N1 blunted affect, N4 social withdrawal, N6 lack of spontaneity and flow of conversation). PANSS-6 is psychometrically valid, reliable, sensitive to change, scalable and can be used to define remission in patients with psychosis.³⁰ Symptomatic remission is defined as a PANSS-6 total score <14 and a score of mild or less (≤ 3) on each of its six items.^{30,33}

The abbreviated version of the YMRS is used to evaluate the six core clinical symptoms of a manic episode: elevated mood (item 1), increased motor activity/energy (item 2), decreased need for sleep (item 4), increased rate or amount of speech (item 6), language/thought disorder (racing thoughts or flight of ideas) (item 7) and insight (excessive involvement in risky activities) (item 11).³¹ The YMRS-6 is a brief instrument that measures treatment response in manic and mixed episodes adequately. Clinical remission is defined as a score of <4.^{31,34}

The MADRS-6 is a six-item subscale score of the original test, widely acknowledged to focus on the core symptoms of depression: apparent sadness (item 1), reported sadness (item 2), inner tension (item 3), lassitude (item 7), inability to feel (item 8) and pessimistic thoughts (item 9).³² The MADRS-6 is a reliable instrument for

assessing antidepressant efficacy in affective disorders. A score of <5 points is considered clinical remission.^{32,34}

Proxy tools for the assessment of the PANSS-6, YMRS-6 and MADRS-6, and methodology for measuring in EHR-based studies

For scoring the PANSS-6, YMRS-6 and MADRS-6 proxies, clinicians carefully read the patient's EHR considering the setting in which the clinical note was written (emergency unit, in-patient admission, community follow-up, etc.). Patients with non-affective psychosis were assessed with the PANSS-6 proxy, and those with affective psychosis were also assessed with the YMRS-6 proxy and/or the MADRS-6 proxy, depending on whether it was a case of mania, depression or mixed episodes with psychotic symptoms. When symptoms were not reported as such in the medical record (relatively commonplace in routine practice because of time constraints during visits), clinicians had to complete the missing data by contextualising them based on the clinical setting (emergency department, in-patient admission, community follow-up, etc.), their experience and knowledge of the psychotic disorder being assessed.

The proxies and their scoring are presented in Fig. 1. Proxies for psychotic and affective features were constructed with a symptom checklist based on items from the PANSS-6, YMRS-6 and MADRS-6. Symptoms were classified as absent/mild, present or severe according to the following criteria: (a) 'absent/mild' when symptoms are absent or at the upper end of normal limits or are mild and intermittent; (b) 'present' when symptoms are relatively stable and cause moderate distress and affect the patient; and (c) 'severe' when symptoms significantly interfere with functioning, interpersonal relationships and/or may seriously compromise the safety of the patient or others. The PANSS-6 proxy scores were 1 (absent/mild), 3 (present) and 6 (severe), with total scores ranging from 6 to 36 points. Clinical remission for the proxy PANSS-6 was defined as a score of 6, with all symptoms classified as absent/mild. The YMRS-6 and MADRS-6 proxies were scored as 0 (absent/mild), 3 (present) and 6 (severe), with total scores ranging from 0 to 36 points. For these proxies, a score of ≤ 3 was defined as clinical remission, with all but one of the symptoms classified as absent or mild. An example of the scoring method is shown in Fig. 2.

Standardised procedure for the validation of PANSS-6, YMRS-6 and MADRS-6 proxy measures

The proxy measures were validated in the context of real-world practice with patients with FEP. Clinicians participating in the study were trained in the use of the assessment tools and followed a manual with instructions designed by the first author. Thus, during a clinical interview, a first clinician recorded a clinical note describing the patient's mental condition, and a second clinician completed the psychometric assessment with the PANSS-6, YMRS-6 and/or MADRS-6 as appropriate, depending on whether the patient had a non-affective or affective psychosis. A third clinician, not present during the clinical interview and blinded to the scale scores, then carefully read the clinical notes in the patient's EHR and examined the severity of symptoms with the proxy tools. The concordance between the scale and proxy scores, using a variety of statistical methods, determined the validity of the proxy instruments.

DIRAYA: the electronic health information system of the Andalusian Health Service

The main cohort's EHR data were collected from DIRAYA, the digital health record software used in the Andalusian public

health system (Spain). The DIRAYA is an EHR system that integrates all of the health information for each patient into a single regional record linking all of the clinical data from primary care centres, specialist out-patient clinics, hospitals, emergency services and pharmacies, including electronic drug prescriptions, telemedicine, referrals, radiology, laboratory tests, appointments, etc. The DIRAYA is one of the largest registration systems in Europe, containing clinical information on the more than 8 million people in the Andalusian public health network, and is therefore a very valuable resource for epidemiological studies.^{35,36}

Data analysis

Descriptive statistics for patient characteristics were expressed as percentages, mean or median, and s.d. or interquartile range (IQR), as appropriate. Normal distribution was tested with the Kolmogorov–Smirnov test. Differences between PANSS-6, YMRS-6 and MADRS-6 and their respective proxy versions were examined by univariate and multivariate analyses. An *a priori* power analysis was performed to calculate the appropriate sample size. With an expected correlation coefficient of 0.6, an assumed type 1 error of 0.05 and a type 2 error rate of 0.20, each of the proxies required a minimum sample size of 19 to achieve adequate statistical power. Either the Pearson's or Spearman's coefficient was used for binary correlations between scores on each instrument, depending on whether or not the normality criteria were met. Partial correlation coefficients were performed to estimate the relationships between the psychometric scales and proxy scores, controlling for potential confounders such as type of psychosis (non-affective or affective), time of assessment (acute or stable) and clinician experience in scoring the proxies (less than or over 5 years). Scatter plots were drawn to display the correlations between variables. The extent to which the proxies predicted clinical remission was determined by regression analysis (using the above cut-off points for each scale and its corresponding proxy). Odds ratios and 95% confidence intervals were estimated by logistic regression. The Nagelkerke R^2 was used to assess the variance explained by the model. A classification table and area under the receiver operating characteristic curve (AUC) were used to evaluate the accuracy and discriminatory ability of the proxies. Internal validation was performed with bootstrapping (2000 resamples). External validation was performed by repeating the same statistical analyses described above in the TIPP-Lausanne cohort patients with early psychosis. Thus, the replicability of the findings in a different data-set would ensure the generalisability of the proxy tools. Interrater reliability of the proxy method was tested by intraclass correlation coefficient (ICC), two-way random effects model and absolute agreement. To this end, one of the participating patients diagnosed with a psychotic episode with mixed affective features was selected at random to assess the interrater reliability of the PANSS-6, YMRS-6 and MADRS-6 proxies, using a single case history. The ratings of this patient's clinical notes were examined by 20 different raters across all of the participating centres for ICC calculation. Missing data were managed with the listwise deletion method. Significance was set at $P < 0.05$. All statistical analyses were conducted with SPSS Statistics version 24 for MacOS (IBM Corporation, Armonk, New York, USA) and MedCalc Statistical Software version 20 for Mac OS (MedCalc Software, Ostend, Belgium; see <https://www.medcalc.org>).

Results

Patient sample characteristics

Table 1 shows the patient sociodemographic and clinical characteristics. From April 2022 to December 2022, 116 patients with FEP

PROXY-BASED ASSESSMENT: EXAMINER'S GUIDE

Assess the symptoms after carefully reading the patient's clinical notes and considering the care setting in which the mental health assessment was conducted (emergency department, in-patient admission, community follow-up, etc.). If any of the symptoms are not reported as such in the medical record, complete the missing data by contextualising them based on the clinical setting and your experience and knowledge of the psychotic disorder being assessed.

Use the following scoring system:

- **Absent/minimal**: symptoms are absent or at the upper end of normal limits or are mild and intermittent
- **Present**: symptoms are relatively stable and cause moderate distress and affect the patient
- **Severe**: symptoms significantly interfere with functioning, interpersonal relationships and/or may seriously compromise the safety of the patient or others.

PROXY ASSESSMENT OF PSYCHOTIC SYMPTOMS (time frame: past week)

- 1) Delusions: Absent/Minimal Present Severe
- 2) Conceptual disorganisation: Absent/Minimal Present Severe
- 3) Hallucinations: Absent/Minimal Present Severe
- 4) Blunted affect: Absent/Minimal Present Severe
- 5) Social withdrawal: Absent/Minimal Present Severe
- 6) Lack of spontaneity and flow of conversation: Absent/Minimal Present Severe

PROXY ASSESSMENT OF MANIC SYMPTOMS (time frame: past 48 h)

- 1) Elevated mood: Absent/Minimal Present Severe
- 2) Increased motor activity/energy: Absent/Minimal Present Severe
- 3) Decreased need for sleep: Absent/Minimal Present Severe
- 4) Increased rate or amount of speech: Absent/Minimal Present Severe
- 5) Language/thought disorder (racing thoughts of flight of ideas): Absent/Minimal Present Severe
- 6) Insight (excessive involvement in risky activities): Absent/Minimal Present Severe

PROXY ASSESSMENT OF DEPRESSIVE SYMPTOMS (time frame: past week)

- 1) Apparent sadness: Absent/Minimal Present Severe
- 2) Reported sadness: Absent/Minimal Present Severe
- 3) Inner tension: Absent/Minimal Present Severe
- 4) Lassitude: Absent/Minimal Present Severe
- 5) Inability to feel: Absent/Minimal Present Severe
- 6) Pessimistic thoughts: Absent/Minimal Present Severe

Fig. 1 The proxies and their scoring criteria.

diagnosed following ICD-10 criteria (codes F1x.4, F1x.5, F1x.7, F20-29, F30.2, F31.2, F31.5, F31.6, F32.3, F33.3 and F53.1) were recruited. Their median age was 26 years (IQR 22–35 years) and 69.8% ($n = 81$) were men. The majority were European White (83.6%, $n = 97$) and unmarried (79.3%, $n = 92$). Of the total sample, 17.2% ($n = 20$) were living alone, 25% ($n = 29$) had a higher education and 38.8% ($n = 45$) were unemployed. The median age at onset of psychosis was 25.5 years (IQR 21–34 years) and the median duration of untreated psychosis was 1 month (IQR 1–5 months). Nineteen patients (16.4%) had a first-degree family history of psychosis. Poor premorbid adjustment was noted in 31.9% ($n = 37$) of the sample, 45.7% ($n = 53$) had past psychiatric records and 57.8% ($n = 67$) had a concurrent history of harmful or hazardous substance use. The most common diagnosis according to ICD-10 criteria was schizophrenia and schizophreniform disorder (31%, $n = 36$) (see Table 1 for further details). Suicidal symptoms were observed in 22.4% ($n = 26$) of the sample. The majority of the patients with FEP (76.7%, $n = 89$) were hospital in-patients at the time of recruitment.

Psychometric assessment of patients

Psychometric characteristics of the sample are summarised in Table 2. A total of 116 assessments of psychotic symptoms (with the PANSS-6 and its proxy), 26 assessments of manic symptoms (with the YMRS-6 and its proxy) and 31 assessments of depressive symptoms (the MADRS-6 and its proxy) were made. The median PANSS-6 score was 15 (IQR 11–20.75) and the PANSS-6 proxy median was 14 (IQR 8–18.75). In terms of remission criteria, 30.2% ($n = 35$) of patients assessed with the PANSS-6 were classified as in symptomatic remission versus the 23.3% ($n = 27$) who met the proxy criteria. The median was 3.5 (IQR 0–20) for the YMRS-6 and 3 (IQR 0–24.75) for its proxy. Twelve patients (46.2%) met YMRS-6 remission criteria, and 14 (53.8%) were classified as in clinical remission by the proxy. The median MADRS-6 score was 9 (IQR 2–19) and the MADRS-6 proxy median score was also 9 (IQR 3–15). The MADRS-6 classified 29% ($n = 9$) of patients as in clinical remission, and 32.3% ($n = 10$) were classified as in clinical remission by the proxy. More than half of the patients (57.8%, $n = 67$) were assessed in the acute phase of psychosis and over three-fourths (76.7%, $n = 89$) had non-affective

EXAMPLE OF CLINICAL RECORD

Male patient, 28 years old, with no history of previous psychiatric admissions or out-patient treatment, who was admitted to the in-patient unit over the weekend because of his public disturbances. As stated in the emergency report, he was in a bar and started preaching a mystical-messianic delusional ideation. Police intervention became necessary when he refused to leave, presenting an episode of agitation. Yesterday in the ward he showed very disorganised behaviour and had to be seen twice by the duty psychiatrist (verbal de-escalation plus oral medication).

Today, during the interview, he was tense, suspicious and uncooperative, saying very cryptically, "everything is written" (while pointing to the ceiling). It was not possible to go deeper into the psychopathological examination, although he seems to have auditory hallucinations, as there were moments during the interview when he mumbled and laughed for no reason. He is unaware of the psychotic disorder, although he accepts medication.

Urine toxicology: positive for cannabis. Haemogram and serum biochemical parameters: no findings. Brain CT scan: pending. Antipsychotic and benzodiazepine doses adjusted. Family coming to the hospital tomorrow to see the patient and provide further information (duration of untreated psychosis, type of onset, stressors, previous psychosocial functioning, etc.).

PROXY ASSESSMENT OF PSYCHOTIC SYMPTOMS (time frame: past week)

- 1) Delusions: Absent/Minimal Present
 2) Conceptual disorganisation: Absent/Minimal Present
 3) Hallucinations: Absent/Minimal Severe
 4) Blunted affect: Present Severe
 5) Social withdrawal: Present Severe
 6) Lack of spontaneity and flow of conversation: Present Severe
- PROXY PANSS-6 SCORE: 18 POINTS**

Fig. 2 Example of proxy scoring method. CT, computed tomography; PANSS-6, Positive and Negative Syndrome Scale.

Table 1 Sociodemographic and clinical characteristics

	Total sample (n = 116)
Sociodemographic characteristics	
Age (years), median (IQR)	26 (22–35)
Gender (male), n (%)	81 (69.8%)
Ethnicity (European White), n (%)	97 (83.6%)
Marital status (unmarried), n (%)	92 (79.3%)
Cohabitation (living alone), n (%)	20 (17.2%)
Education (higher education), n (%)	29 (25%)
Occupation (unemployed), n (%)	45 (38.8%)
Clinical characteristics	
Age at psychotic onset (years), median (IQR)	25.5 (21–34)
DUP (months), median (IQR)	1 (1–5)
Family history of psychosis (yes), n (%)	19 (16.4%)
Premorbid psychosocial adjustment (poor premorbid adjustment ^a), n (%)	37 (31.9%)
Past psychiatric record (yes), n (%)	53 (45.7%)
History of substance use, n (%)	67 (57.8%)
Psychotic disorder (ICD-10 codes), n (%)	
Schizophrenia and schizophreniform disorder (F20)	36 (31%)
Acute and transient psychotic disorder (F23)	20 (17.2%)
Mania with psychotic symptoms (F30.2)	12 (10.3%)
Substance-induced psychotic disorder (F1x.5)	12 (10.3%)
Other and unspecified psychotic disorders (F28–29)	12 (10.3%)
Depressive disorder with psychotic features (F32.3, F33.3)	9 (7.8%)
Bipolar disorder with psychotic features (F31.2, F31.5, F31.64)	6 (5.2%)
Persistent delusional disorder (F22)	4 (3.4%)
Schizotypal disorder (F21)	3 (2.6%)
Schizoaffective disorder (F25)	2 (1.7%)
Suicidal symptoms, n (%)	26 (22.4%)
In-patient status at the time of recruitment, n (%)	89 (76.7%)

IQR, interquartile range; DUP, duration of untreated psychosis.
 a. The sum of the following sociodemographic variables was used as an indicator of poor premorbid adjustment: being unmarried, unemployed and low education.

psychotic disorders. The majority of the proxies (61.2%, $n = 71$) were completed by experienced clinicians (over 5 years).

Correlation and regression analyses between scales and proxies

Scatter plots and correlation analyses are shown in Fig. 3. The Spearman correlation coefficient (ρ) between the PANSS-6 and its proxy was 0.88 (95% CI 0.83–0.91; $P < 0.001$). After controlling for type of psychosis (non-affective or affective), time of assessment (acute or stable) and clinician experience (less than or over 5 years), the partial rank correlation coefficient was 0.75 (95% CI 0.66–0.82; $P < 0.001$), indicating a strong positive linear relationship between the two measures. Regression analysis showed that the proxy PANSS-6 very strongly predicted psychotic remission (odds ratio 231.11, 95%

Table 2 Psychometric characteristics

Psychometric assessments	
PANSS-6 score ($n = 116$), median (IQR)	15 (11–20.75)
PANSS-6 remission criteria, ^a n (%)	35 (30.2%)
YMRS-6 score ($n = 26$), median (IQR)	3.5 (0–20)
YMRS-6 remission criteria, ^b n (%)	12 (46.2%)
MADRS-6 score ($n = 31$), median (IQR)	9 (2–19)
MADRS-6 remission criteria, ^c n (%)	9 (29%)
Proxy-based assessments	
Proxy PANSS-6 score ($n = 116$), median (IQR)	14 (8–18.75)
Proxy PANSS-6 remission criteria, ^a n (%)	27 (23.3%)
Proxy YMRS-6 score ($n = 26$), median (IQR)	3 (0–24.75)
Proxy YMRS-6 remission criteria, ^b n (%)	14 (53.8%)
Proxy MADRS-6 score ($n = 31$), median (IQR)	9 (3–15)
Proxy MADRS-6 remission criteria, ^c n (%)	10 (32.3%)

PANSS-6, Positive and Negative Syndrome Scale; IQR, interquartile range; YMRS-6, Young Mania Rating Scale; MADRS-6, Montgomery–Åsberg Depression Rating Scale.
 a. Symptomatic remission was defined as a PANSS-6 score < 14 (with a score of ≤ 3 on each of the items) and a proxy PANSS-6 score of 6.
 b. Symptomatic remission was defined as a YMRS-6 score < 4 and a proxy YMRS-6 score ≤ 3 .
 c. Symptomatic remission was defined as a MADRS-6 score < 5 and a proxy MADRS-6 score ≤ 3 .

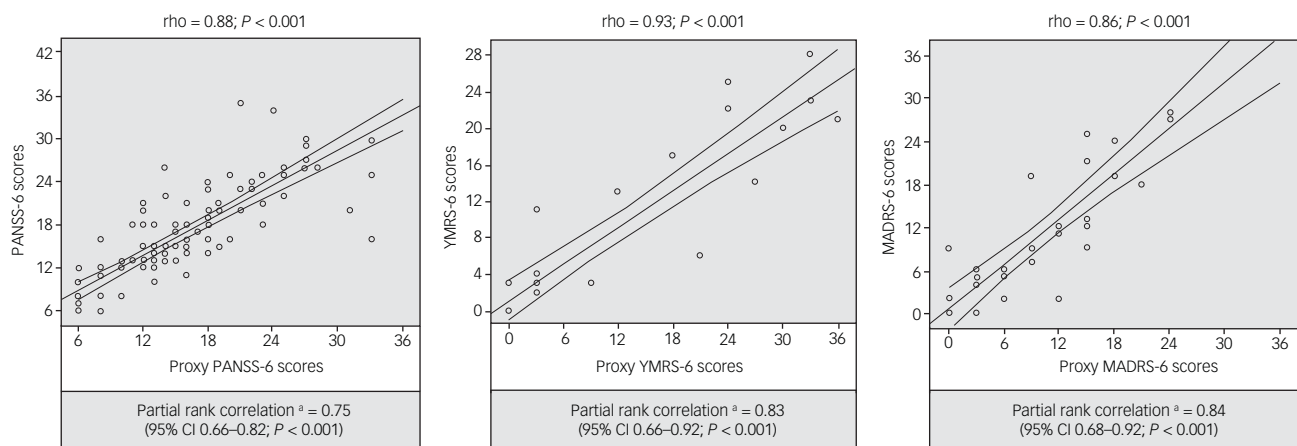


Fig. 3 Scatter plots and correlation analyses between scale scores and their respective proxy scores. ^aAfter controlling for type of psychosis (non-affective or affective), time of assessment (acute or stable) and clinician experience (<5 or 5 years). MADRS-6, Montgomery-Åsberg Depression Rating Scale; PANSS-6, Positive and Negative Syndrome Scale; YMRS-6, Young Mania Rating Scale.

CI 27.94–1911.66; $P < 0.001$). The Nagelkerke's R^2 was 0.68, which means that the model accounted for 68% of the variance. The classification table showed that 91.4% of patients with FEP were classified correctly as in or not in remission. The AUC was 0.93 (95% CI 0.87–0.99) and the bootstrapped AUC was 0.89 (95% CI 0.83–0.95), demonstrating that the model had low overfitting and that the proxy PANSS-6 had excellent predictive power. For the YMRS-6 and its proxy, the Spearman's rho coefficient was 0.93 (95% CI 0.85–0.97; $P < 0.001$), indicating a strong positive correlation between these two variables. The partial rank correlation coefficient was 0.83 (95% CI 0.66–0.92; $P < 0.001$), showing that this strong relationship persisted after adjustment for potential confounders (type of psychosis, time of evaluation and clinician experience in scoring the proxies). Logistic regression showed that the proxy YMRS-6 predicted the psychometric criteria for remission of mania well (odds ratio 66, 95% CI 5.23–833.56; $P = 0.001$), accounting for 66% (Nagelkerke's R^2) of the variance. The model showed that the proxy version correctly classified 88.5% of patients as in or not in remission according to YMRS-6 criteria. The AUC was 0.89 (95% CI 0.74–1) and the bias-corrected AUC after bootstrapping was 0.88 (95% CI 0.70–1), indicating that the model had minimal overfitting and that the YMRS-6 proxy had excellent discriminatory performance. For depressive symptoms, the Spearman coefficient for the correlation between the MADRS-6 and the proxy MADRS-6 was 0.86 (95% CI 0.72–0.93; $P < 0.001$). This strong positive linear relationship remained significant after adjustment for the covariates listed above, with a partial rank coefficient of 0.84 (95% CI 0.68–0.92; $P < 0.001$). Regression analysis confirmed this strong relationship, as the proxy MADRS-6 predicted the clinical remission of depression well (odds ratio 22.17, 95% CI 3.04–161.84; $P = 0.002$). Nagelkerke's R^2 was 0.46. The classification table showed that the MADRS-6 proxy correctly distinguished 83.9% of patients in and not in remission according to the MADRS-6 remission criteria. The AUC was 0.80 (95% CI 0.62–0.99) and the bootstrapped optimism-corrected AUC was 0.76 (95% CI 0.52–1). This indicated low overfitting of the model and good predictive power of the proxy version of the MADRS-6.

External validation and interrater reliability of the proxy tools

A total of 30 assessments of positive, manic and depressive symptoms were made of patients from the TIPP-Lausanne cohort to provide external validation of the above findings. More than half of the

proxies were assessed during the stabilisation phase and in cases of non-affective psychotic disorders (70%, $n = 21$). The majority were completed by clinicians with less than 5 years of experience (66.6%, $n = 20$). The sociodemographic, clinical and psychometric characteristics of these patients are summarised in Supplementary Table 2, whereas scatter plots and correlation analyses are shown in Supplementary Fig. 1 in Supplementary Appendix 1. The Spearman correlation coefficient between the PANSS-6 and the proxy was 0.82 (95% CI 0.65–0.91; $P < 0.001$), whereas the partial rank correlation coefficient after adjustment for potential confounders (type of psychosis, time of assessment and clinician experience in scoring the proxies) was 0.66 (95% CI 0.39–0.82; $P < 0.001$). Regression analysis showed that the proxy PANSS-6 significantly predicted the psychometric criteria for PANSS-6 remission (odds ratio 20, 95% CI 2.04–196.37; $P = 0.010$). The Nagelkerke's R^2 was 0.40, the model correctly classified 76.7% of patients in and not in remission, and the AUC was 0.75 (95% CI 0.57–0.94). The Spearman coefficient for the correlation between the YMRS-6 and its proxy was 0.76 (95% CI 0.54–0.88; $P < 0.001$), and this strong positive relationship remained significant after adjusting for the above-mentioned covariates, with a partial rank coefficient of 0.77 (95% CI 0.57–0.88; $P < 0.001$). Regression analysis showed that the proxy YMRS-6 significantly predicted the psychometric criteria for remission from mania (odds ratio 26, 95% CI 3.69–183.42; $P = 0.001$), with a Nagelkerke's R^2 for the model of 0.52. It correctly classified 83.3% of patients in and out of remission, and the AUC was 0.84 (95% CI 0.68–0.99). The Spearman coefficient for the correlation between the MADRS-6 and its proxy was 0.83 (95% CI 0.66–0.91; $P < 0.001$), whereas the partial rank correlation coefficient, after controlling for potential confounders, was 0.84 (95% CI 0.69–0.92; $P < 0.001$). Regression analysis showed that the proxy version significantly predicted MADRS-6 remission criteria (odds ratio 28, 95% CI 2.82–277.96; $P = 0.004$). Nagelkerke's R^2 was 0.47, the model correctly discriminated 80% of patients in and not in remission according to MADRS-6 criteria, and its AUC was 0.82 (95% CI 0.67–0.98). Finally, the overall ICC of the proxies calculated to quantify agreement between raters was 0.81 (95% CI 0.52–0.99; $P < 0.001$), indicating that the proxy method had good interrater reliability.

Discussion

The aim of this study was to develop and examine the validity and reliability of PANSS-6, YMRS-6 and MADRS-6 proxy measures

based on information in clinical records from EHRs in a real-world setting. The main findings were that the proxy versions were strongly correlated with their respective standardised scales and had good accuracy and discriminatory power in distinguishing between patients in and not in remission, with high interrater reliability. These findings were not altered when the effects of other explanatory variables, such as type of psychosis (non-affective or affective), time of assessment (acute or stabilisation phase) and clinician experience (less than or over 5 years), were controlled for. The proxies were also tested in an independent FEP cohort, showing good external validation, which suggests that the proxy method for assessing psychotic, manic and depressive symptoms could be generalised to any other EHR data-set. To the best of our knowledge, this is the first study validating EHR-based proxy tools for the psychometric assessment of psychotic and affective symptoms in patients with FEP under routine clinical practice conditions.

As there are no other validated clinician-administered proxy reporting instruments for the assessment of psychotic and affective symptoms in patients with FEP, the results of the development and validation of our proxy tools had to be compared with other available validated, clinician-rated proxy measures in the field of psychosis. Thus, the correlation coefficients for the proxy PANSS-6, YMRS-6 and MADRS-6, which ranged from 0.75 to 0.93, were considerably higher than the 0.57 found in validating a proxy measure of the Premorbid Adjustment Scale (PAS).²⁰ Accuracy of remission status classification by the PANSS-6, YMRS-6 and MADRS-6 proxies (83.9–91.4%) was similar to the 82% achieved with the PAS proxy and the 87% achieved with the proxy case identification tool for deficit schizophrenia (PDS).^{17,20} Finally, AUCs indicating the discriminative power of the PANSS-6, YMRS-6 and MADRS-6 proxies ranged from 0.80 to 0.93, which was slightly better than the 0.78 achieved by the PAS proxy and the 0.73 achieved by a proxy measure of PANSS remission criteria based on a CGI-Improvement Scale value of 1.^{20,37} The interrater reliability of our proxy methodology (ICC = 0.81) was similar to that found by Fenton and McGlashan,³⁸ when they retrospectively scored the Schedule for the Assessment of Positive Symptoms (SAPS) and the Schedule for the Assessment of Negative Symptoms (SANS) by reviewing the admission records of patients with FEP. Although those authors did not validate their methodology, they found high interrater agreement of the SAPS (ICC = 0.85) and SANS (ICC = 0.83) proxies from manual chart review.³⁸ All of the above comparisons suggest that our proxy versions of the PANSS-6, YMRS-6 and MADRS-6 meet validity and reliability standards, and can be used to measure the severity of psychotic and affective symptoms in EHRs.

Methodological limitations

Our findings should be viewed in the light of the following three main limitations. First, there were sample size-related limitations, particularly in the validation of the YMRS and MADRS proxies, which yielded some very high correlations in the univariate analysis that exceeded the reliability of the scales. This phenomenon can occur by chance and the real value, although probably still very high, is likely to be lower.³⁹ Notwithstanding, our sample size was still larger than the minimum required to achieve statistical power for correlation analyses. Second, there were limitations arising from the simplicity of the proxies, in that although they assess the severity of the core features of psychotic, manic and depressive symptoms, they do not capture other psychopathological features with significant predictive value in patients with psychosis, such as psychomotor disturbances or first-rank symptoms of schizophrenia.⁴⁰ Third, despite the simplicity of these proxy instruments, another limitation of the methodology would be the

time required to accurately assess symptom severity and estimate missing data by clinical contextualisation through careful reading of patients' clinical notes. Fourth, the quality of the clinical notes in the primary cohort may have been somewhat biased because clinicians who wrote them knew that their notes would later be used to construct a proxy measure, and this could have influenced how they recorded their findings on patients' mental states. In this regard, we would like to emphasise that the EHR clinical notes on which the proxy ratings were based were similar to those routinely collected in real-world practice, and that this potential bias was not present in the external validation, as the notes on the TIPP-Lausanne cohort were written before it joined this study. Finally, there is the inherent limitation of a cross-sectional design in capturing changes in patients' psychopathology over time. However, it is worth noting that our proxy methodology discriminates well between patients in and not in remission, and that the time of psychopathological assessment (acute or stabilisation phase) was a covariate in the multivariate analyses.

Clinical implications and future directions

Despite these limitations, our proxies for the abbreviated PANSS, YMRS and MADRS have meaningful implications. They are instruments based on real-life settings valuable for large-scale collaborative research, as they could make different databases psychometrically compatible. They could also be useful tools for homogenising clinical information in case report-based systematic reviews. Beyond the research framework, they would also be useful for monitoring the progress of patients with severe mental disorders in clinical practice, and for examining the effectiveness of mental health treatment programmes. Future directions in this area should study whether these proxy measures can be administered by other healthcare professionals (such as general practitioners, nurses or social workers), or whether they would be valid for use with children and adolescents. Similarly, as proxies were validated in high-income countries with well-structured, resourced mental health services, they should also be validated in low- and middle-income countries with fewer mental health resources, where not everyone with FEP has access to early intervention programmes. Finally, it is important to highlight the opportunities that artificial intelligence and natural language processing would offer in analysing the unstructured free text in the EHRs to automatically provide proxy estimates of patients' clinical status. In this respect, it would be of great interest to integrate our proxy measures with SAVANA's EHRRead and CRIS-CODE technology.^{41,42} Such integration with artificial intelligence with natural language processing algorithms could somewhat alleviate the ongoing ethical dilemma in biomedical EHR-based research regarding the extraction of data from patients' medical records without informed consent.^{43,44}

In summary, we conclude that standardised psychometric assessments are essential for patients with FEP. Therefore, our proposal of a proxy instrument based on clinical records for the assessment of psychotic and affective symptoms offers a feasible, valid and reliable alternative to standard psychometric procedures in real-world practice settings. This would greatly expand opportunities for observational studies using routinely collected health data, and would be particularly useful in research-based on EHRs.

Álvaro López-Díaz , Mental Health Clinical Management Unit, Virgen Macarena University Hospital, Seville, Spain; Translational Psychiatry Research Group (PsyNal), Seville Biomedical Research Centre (IBIS), Spain; Spanish Network for Research in Mental Health, Carlos III Institute (CIBERSAM, ISCIII), Seville, Spain; Department of Psychiatry, School of Medicine, University of Seville, Spain; and First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; **Fernanda Jazmín Palermo-Zeballos**, Mental Health Clinical Management Unit, Virgen Macarena University Hospital, Seville, Spain; and First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; **Luis Gutiérrez-Rojas**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; Mental Health Clinical Management Unit, San Cecilio University Hospital, Granada, Spain; Psychiatry and Neurosciences Research Group (CTS-549), Institute of

Neurosciences, University of Granada, Spain; and Department of Psychiatry, University of Granada, Spain; **Luis Alameda**, Service of General Psychiatry, Treatment and Early Intervention in Psychosis Program (TIPP), Lausanne University Hospital and University of Lausanne, Switzerland; and Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK; **Francisco Gotor-Sánchez-Luengo**, Department of Psychiatry, School of Medicine, University of Seville, Spain; First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Virgen del Rocío University Hospital, Seville, Spain; **Nathalia Garrido-Torres**, Translational Psychiatry Research Group (PsyNal), Seville Biomedical Research Centre (IBiS), Spain; Spanish Network for Research in Mental Health, Carlos III Institute (CIBERSAM, ISCIII), Seville, Spain; First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Virgen del Rocío University Hospital, Seville, Spain; **Johann Métrailler**, Service of General Psychiatry, Treatment and Early Intervention in Psychosis Program (TIPP), Lausanne University Hospital and University of Lausanne, Switzerland; **Livia Alerci**, Service of General Psychiatry, Treatment and Early Intervention in Psychosis Program (TIPP), Lausanne University Hospital and University of Lausanne, Switzerland; **Vincent Bonnairel**, Service of General Psychiatry, Treatment and Early Intervention in Psychosis Program (TIPP), Lausanne University Hospital and University of Lausanne, Switzerland; **Pablo Cano-Domínguez**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Virgen de la Victoria University Hospital, Málaga, Spain; **Elma Avanesi-Molina**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Virgen de la Victoria University Hospital, Málaga, Spain; **Miguel Soto-Ontoso**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Torrecárdenas University Hospital, Almería, Spain; **Rocío Torrecilla-Olavarrieta**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Jerez University Hospital, Cádiz, Spain; **Leticia Irene Muñoz-Manchado**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Jerez University Hospital, Cádiz, Spain; **Pedro Torres-Hernández**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Jaén University Hospital, Spain; **Fermin González-Higueras**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Jaén University Hospital, Spain; **Juan Luis Prados-Ojeda**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Reina Sofía University Hospital, Córdoba, Spain; **Mario Herrera-Cortés**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Reina Sofía University Hospital, Córdoba, Spain; **José Miguel Meca-García**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Poniente University Hospital, Almería, Spain; **Rafael Manuel Gordillo-Urbano**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Infanta Margarita Hospital, Córdoba, Spain; **Cristina Sánchez-Robles**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Juan Ramón Jiménez Hospital, Huelva, Spain; **Tomás Delgado-Durán**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Juan Ramón Jiménez Hospital, Huelva, Spain; **María Felipa Soriano-Peña**, First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, San Agustín University Hospital, Linares, Spain; **Philippe Golay**, Service of General Psychiatry, Treatment and Early Intervention in Psychosis Program (TIPP), Lausanne University Hospital and University of Lausanne, Switzerland; **Philippe Conus**, Service of General Psychiatry, Treatment and Early Intervention in Psychosis Program (TIPP), Lausanne University Hospital and University of Lausanne, Switzerland; **Benedicto Crespo-Facorro**, Translational Psychiatry Research Group (PsyNal), Seville Biomedical Research Centre (IBiS), Spain; Spanish Network for Research in Mental Health, Carlos III Institute (CIBERSAM, ISCIII), Seville, Spain; Department of Psychiatry, School of Medicine, University of Seville, Spain; First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Virgen del Rocío University Hospital, Seville, Spain; **Miguel Ruiz-Veguilla**, Translational Psychiatry Research Group (PsyNal), Seville Biomedical Research Centre (IBiS), Spain; Spanish Network for Research in Mental Health, Carlos III Institute (CIBERSAM, ISCIII), Seville, Spain; Department of Psychiatry, School of Medicine, University of Seville, Spain; First-Episode Psychosis Research Network of Andalusia (Red PEPsur), Spain; and Mental Health Clinical Management Unit, Virgen del Rocío University Hospital, Seville, Spain

Correspondence: Álvaro López-Díaz. Email: alvaro.lopez.diaz.sspa@jurtadeandalucia.es

First received 24 Aug 2023, final revision 31 Oct 2023, accepted 14 Nov 2023

Supplementary material

Supplementary material is available online at <https://doi.org/10.1192/bjo.2023.623>

Data availability

The data that support the findings of this study are available from the corresponding author, A.L.-D., upon reasonable request.

Author contributions

A.L.-D. devised the study concept and design, collected data, coordinated the data collection, managed the statistical analyses, interpreted the results, wrote the first draft of the manuscript and edited the final version. L.G.-R. collected data, contributed to drafting the manuscript and approved its final version. J.M., L. Alerci and V.B. collected data, reviewed the contents of the manuscript and approved the final version. L. Alameda, P.G., P.C., M.R.-V. and B.C.-F. critically revised the manuscript and approved the final version. All other authors collected data, read the final manuscript and approved it for publication.

Funding

M.R.-V. is supported by grants from the Alicia Koplowitz Foundation and the project P122/00678, funded by Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union. L. Alameda is supported by a fellowship of the Adrian and Simone Frutiger Foundation and by the Carigest Foundation.

Declaration of interest

None.

References

- 1 Fusar-Poli P, McGorry PD, Kane JM. Improving outcomes of first-episode psychosis: an overview. *World Psychiatry* 2017; **16**(3): 251–65.
- 2 Ramin J, Conus P, Golay P. Exploring the clinical relevance of a dichotomy between affective and non-affective psychosis: results from a first-episode psychosis cohort study. *Early Interv Psychiatry* 2022; **16**(2): 168–77.
- 3 Fusar-Poli P, Cappucciati M, Rutigliano G, Heslin M, Stahl D, Brittenden Z, et al. Diagnostic stability of ICD/DSM first episode psychosis diagnoses: meta-analysis. *Schizophr Bull* 2016; **42**(6): 1395–406.
- 4 Baer L, Blais MA. *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health*. Humana Press, 2010.
- 5 Kay SR, Fiszbein A, Opfer LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; **13**(2): 261–76.
- 6 Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychol Rep* 1962; **10**: 799–812.
- 7 Haro JM, Kamath SA, Ochoa S, Novick D, Rele K, Fargas A, 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; **107**(416): 16–23.
- 8 Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978; **133**: 429–35.
- 9 Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134**(4): 382–9.
- 10 Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23**(1): 56–62.
- 11 Østergaard SD, Opler MGA, Correll CU. Bridging the measurement gap between research and clinical care in schizophrenia: Positive and Negative Syndrome Scale-6 (PANSS-6) and other assessments based on the Simplified Negative and Positive Symptoms Interview (SNAPSI). *Innov Clin Neurosci* 2017; **14**(11): 68–72.
- 12 Golay P, Baumann PS, Jenni R, Do KQ, Conus P. Patients participating to neurobiological research in early psychosis: a selected subgroup? *Schizophr Res* 2018; **201**: 249–53.
- 13 Hrisos S, Eccles MP, Francis JJ, Dickinson HO, Kaner EF, Beyer F, et al. Are there valid proxy measures of clinical behaviour? A systematic review. *Implement Sci* 2009; **4**(1): 37.
- 14 Torres Soler C, Olofsson S, Vadlin S, Ramklint M, Nilsson KW, Sonnyby K. Diagnostic accuracy of the Montgomery-Åsberg Depression Rating Scale parent report among adolescent psychiatric outpatients. *Nord J Psychiatry* 2018; **72**(3): 184–90.
- 15 Gracious BL, Youngstrom EA, Findling RL, Calabrese JR. Discriminative validity of a parent version of the Young Mania Rating Scale. *J Am Acad Child Adolesc Psychiatry* 2002; **41**(11): 1350–9.
- 16 Crawford JR, Parker DM, Stewart LE, Besson JAO, De Lacey G. Prediction of WAIS IQ with the National Adult Reading Test: cross-validation and extension. *Br J Clin Psychol* 1989; **28**(3): 267–73.
- 17 Kirkpatrick B, Buchanan RW, Breier A, Carpenter WT. Case identification and stability of the deficit syndrome of schizophrenia. *Psychiatry Res* 1993; **47**(1): 47–56.
- 18 Drukker M, Krabbendam L, Driessen G, van Os J. Social disadvantage and schizophrenia: a combined neighbourhood and individual-level analysis. *Soc Psychiatry Psychiatr Epidemiol* 2006; **41**(8): 595–604.
- 19 Monk-cunliffe J, Kadra-scalzo G, Finamore C, Dale O, Khondoker M, Barrett B, et al. Defining severity of personality disorder using electronic health records: short report. *BJPsych Open* 2023; **9**(5): e137.
- 20 López-Díaz Á, Ayesa-Arriola R, Garrido-Torres N, Otíz-García de la Foz V, Suárez-Pinilla P, Ramírez-Bonilla ML, et al. A proxy measure of pre-morbid adjustment in psychosis for large-scale epidemiological studies and electronic health record-based research. *Schizophr Res* 2022; **243**: 467–9.
- 21 Baumann PS, Crespi S, Marion-Veyron R, Solida A, Thonney J, Favrod J, et al. Treatment and Early Intervention in Psychosis Program (TIPP-Lausanne):

- implementation of an early intervention programme for psychosis in Switzerland. *Early Interv Psychiatry* 2013; **7**(3): 322–8.
- 22 Sheehan D V, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, 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): 22–33.
 - 23 Yung AR, Yuen HP, McGorry PD, Phillips LJ, Kelly D, Dell’Olio M, et al. Mapping the onset of psychosis: the comprehensive assessment of at-risk mental states. *Aust N Z J Psychiatry* 2005; **39**(11–12): 964–71.
 - 24 World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. World Health Organization, 1993.
 - 25 Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007; **370**: 1453–7.
 - 26 Singh SP, Cooper JE, Fisher HL, Tarrant CJ, Lloyd T, Banjo J, et al. Determining the chronology and components of psychosis onset: the Nottingham Onset Schedule (NOS). *Schizophr Res* 2005; **80**(1): 117–30.
 - 27 Silverman MM, Berman AL, Sanddal ND, O’Carroll PW, Joiner TE. Rebuilding the tower of babel: a revised nomenclature for the study of suicide and suicidal behaviors part 1: background, rationale, and methodology. *Suicide Life Threat Behav* 2007; **37**(3): 248–63.
 - 28 Silverman MM, Berman AL, Sanddal ND, O’Carroll PW, Joiner TE. Rebuilding the tower of babel: a revised nomenclature for the study of suicide and suicidal behaviors part 2: suicide-related ideations, communications, and behaviors. *Suicide Life Threat Behav* 2007; **37**(3): 264–77.
 - 29 Paykel ES, Myers JK, Lindenthal JJ, Tanner J. Suicidal feelings in the general population: a prevalence study. *Br J Psychiatry* 1974; **124**(5): 460–9.
 - 30 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): 436–44.
 - 31 Tohen M, Kanba S, McIntyre RS, Fujikoshi S, Katagiri H. Efficacy of olanzapine monotherapy in the treatment of bipolar depression with mixed features. *J Affect Disord* 2014; **164**: 57–62.
 - 32 Tohen M, Katagiri H, Fujikoshi S, Kanba S. Efficacy of olanzapine monotherapy in acute bipolar depression: a pooled analysis of controlled studies. *J Affect Disord* 2013; **149**(1–3): 196–201.
 - 33 Andreasen NC, Carpenter WTJ, Kane JM, Lasser RA, Marder SR, Weinberger DR. Remission in schizophrenia: proposed criteria and rationale for consensus. *Am J Psychiatry* 2005; **162**(3): 441–9.
 - 34 Berk M, Ng F, Wang WV, Calabrese JR, Mitchell PB, Malhi GS, et al. The empirical redefinition of the psychometric criteria for remission in bipolar disorder. *J Affect Disord* 2008; **106**(1–2): 153–8.
 - 35 Sánchez-Teruel D, Robles-Bello MA, Sarhani-Robles A. Suicidal vulnerability in older adults and the elderly: study based on risk variables. *BJPsych Open* 2022; **8**(3): e77.
 - 36 Protti D. Moving toward a single comprehensive electronic health record for every citizen in Andalucía, Spain. *Healthc Q* 2007; **10**(4): 114–23.
 - 37 Masand P, O’Gorman C, Mandel FS. Clinical Global Impression of Improvement (CGI-I) as a valid proxy measure for remission in schizophrenia: analyses of ziprasidone clinical study data. *Schizophr Res* 2011; **126**(1–3): 174–83.
 - 38 Fenton WS, McGlashan TH. Natural history of schizophrenia subtypes. *Arch Gen Psychiatry* 1991; **48**(11): 978.
 - 39 Fiedler K. Voodoo correlations are everywhere-not only in neuroscience. *Perspect Psychol Sci* 2011; **6**(2): 163–71.
 - 40 López-Díaz Á, Fernández-González JL, Lara I, Crespo-Facorro B, Ruiz-Veguilla M. The prognostic role of catatonia, hallucinations, and symptoms of schizophrenia in acute and transient psychosis. *Acta Psychiatr Scand* 2019; **140**(6): 574–85.
 - 41 Hernandez Medrano I, Tello Guijarro J, Belda C, Urena A, Salcedo I, Espinosa-Anke L, et al. Savana: re-using electronic health records with artificial intelligence. *Int J Interact Multimed Artif Intell* 2018; **4**(7): 8.
 - 42 Jackson RG, Patel R, Jayatileke N, Kolliakou A, Ball M, Gorrell G, et al. Natural language processing to extract symptoms of severe mental illness from clinical text: the clinical record interactive search comprehensive data extraction (CRIS-CODE) project. *BMJ Open* 2017; **7**(1): e012012.
 - 43 Laurijssen SJ, van der Graaf R, van Dijk WB, Schuit E, Groenwold RH, Grobbee DE, et al. When is it impractical to ask informed consent? A systematic review. *Clin Trials* 2022; **19**(5): 545–60.
 - 44 Piasecki J, Walkiewicz-Żarek E, Figas-Skrzypulec J, Kordecka A, Dranseika V. Ethical issues in biomedical research using electronic health records: a systematic review. *Med Health Care Philos* 2021; **24**(4): 633–58.

