Developing an International Standard Set of Patient-Reported Outcome Measures for Psychotic Disorders

Emily McKenzie, M.Sc., Lucy Matkin, M.Sc., Luz Sousa Fialho, D.Phil., Ifeoma Nneka Emelurumonye, M.D., Timea Gintner, M.A., Christiana Ilesanmi, B.Sc., Beth Jagger, B.A., Shannon Quinney, B.Sc., Elizabeth Anderson, B.C.C., Lone Baandrup, M.D., Ph.D., Amrit Kumar Bakhshy, Alison Brabban, D.Clin.Psy., Tim Coombs, M.M., Ph.D., Christoph U. Correll, M.D., Caroline Cupitt, M.Sc., P.G.Dip., Anju Devianee Keetharuth, Ph.D., Dania Nimbe Lima, Ph.D., Paul McCrone, Ph.D., Mary Moller, Ph.D., D.N.P., Cornelis L. Mulder, Ph.D., David Roe, Ph.D., Grant Sara, M.B.B.S., Ph.D., Farhad Shokraneh, Ph.D., Jacqueline Sin, Ph.D., R.M.N., Kristen A. Woodberry, M.S.W., Ph.D., Donald Addington, M.B.B.S., M.R.C.Psych.; for the Psychotic Disorders Working Group of the International Consortium for Health Outcomes Measurement

Objective: The objective of this project was to develop a set of patient-reported outcome measures for adolescents and adults who meet criteria for a psychotic disorder.

Methods: A research team and an international consensus working group, including service users, clinicians, and researchers, worked together in an iterative process by using a modified Delphi consensus technique that included videoconferencing calls, online surveys, and focus groups. The research team conducted systematic literature searches to identify outcomes, outcome measures, and risk adjustment factors. After identifying outcomes important to service users, the consensus working group selected outcome measures, risk adjustment factors, and the final set of outcome measures. International stakeholder groups consisting of >100 professionals and service users reviewed and commented on the final set.

Results: The consensus working group identified four outcome domains: symptoms, recovery, functioning, and treatment. The domains encompassed 14 outcomes of importance to service users. The research team identified 131 measures from the literature. The consensus working group selected nine measures in an outcome set that takes approximately 35 minutes to complete.

Conclusions: A set of patient-reported outcome measures for use in routine clinical practice was identified. The set is free to service users, is available in at least two languages, and reflects outcomes important to users. Clinicians can use the set to improve clinical decision making, and administrators and researchers can use it to learn from comparing program outcomes.

Psychiatric Services 2022; 73:249–258; doi: 10.1176/appi.ps.202000888

A patient-reported outcome (PRO), as measured by patient-reported outcome measures (PROMs), is any aspect of a patient's health status that comes directly from the patient (1). PROMs can be used to improve clinical care (2, 3); to inform clinicians and patients about treatment progress; to create, compare, and aggregate scores at a high level to inform policy; and to inform approval of drugs and devices (4, 5). Research and application of PROMs in health care, particularly in the management of chronic disorders, have increased over the past 20 years (6). The use of PROMs is a focus of patient-centered care (7). Notwithstanding the challenges to implementing large-scale PROMs systems in health care, factors that increase the rate of PROM collection include provider training in the use of PROMs, use of software to register and work with PROMs in daily practice, administrative surveillance of collection rates, and the presence of local

HIGHLIGHTS

- Patient-reported outcomes reflect the outcomes important to patients, and measures of these outcomes can be used for comparisons across programs and countries.
- An international working group used systematic searches to identify and assess the quality of psychometric evidence supporting patient-reported outcome measures for psychosis and associated risk adjustment factors.
- Service users, clinicians, and researchers were involved in a consensus process to reduce the measures to a standard set that can be completed in about 35 minutes.
- The patient-reported outcomes are assessed by nine measures that cover symptoms, recovery, functioning, and treatment.

clinical champions (3). International application of PROMs requires high standards for translation (8).

We identified two examples of large-scale implementation of recommended PROMs to mental health services in specific programs in the United Kingdom (9) and in routine outcome measurement in Australia (10). However, a Cochrane review of routine monitoring that used PROMs for improving treatment of common mental disorders among adults found insufficient evidence to support routine monitoring and identified the need for "more research of better quality," including measuring a range of relevant outcomes (11).

The outcomes important to patients can be identified through focus groups, in-depth interviews, target population surveys, qualitative synthesis of the literature, and content analysis of available data sources (12). Ideally, PROMs display strong psychometric properties, including a conceptual and measurement model, reliability, validity, responsiveness, interpretability, alternative modes of administration, and cross-cultural and linguistic adaptations (13, 14). Practical implementation of measures warrants consideration and includes identification of the goals for collecting PROs, selection of patients, and determination of the setting and timing of assessments (15).

Psychotic disorders, including schizophrenia and bipolar disorder, represent significant burdens to service users, families, and health systems worldwide (16). Although schizophrenia is a low-prevalence disorder, with an estimated population prevalence of <1% (17), it is associated with adverse mental and general medical health outcomes, a high degree of disability, high health care costs, and a 15-year reduction in life expectancy (18–23). Bipolar disorder has a slightly higher prevalence, at approximately 1%, with psychiatric as well as general medical burden resulting from its early onset, severity, and chronicity (21–25). Although clinical recovery is defined as the remission of symptoms and the return of functioning (26), the meaning of personal recovery to service users is broader and a process that encompasses many aspects of life and promotion of an individual's strength and potential (27). Evidence supports recovery as a feasible outcome for schizophrenia spectrum disorders and bipolar disorder (28, 29). At the health systems level, there has been a shift in focus toward a recovery orientation and personalized care (30, 31). The broad impact of psychotic disorders has spurred investigators to examine a range of PROs in schizophrenia and patient-reported quality of life (32) and functional outcomes (33) in bipolar disorder.

Through use of a consensus-building process, the International Consortium for Health Outcomes Measurement (ICHOM) developed and implemented standard sets of PROMs for use in routine clinical practice for various medical conditions. The process is supported by identification of the evidence and by systematic, critical evaluation of measures and their psychometric properties. ICHOM organized a working group of psychosis experts, including clinicians, researchers, and service users, to identify a set of PROMs to monitor individual treatment outcomes or to compare outcomes of similar mental health services, with a view to establishing the value of these services. The value of health care can be defined as the patient outcomes relative to the costs for obtaining those outcomes (34). Outcome assessment can be guided by a set of PROs for a specific disorder, as exemplified by the set for depression and anxiety developed by ICHOM (35). The specific aims of this study were to develop a standard minimum set of PROMs for psychotic disorders that can be used anywhere in the world and to identify a set of risk adjustment factors to enhance utility of comparisons across treatment modalities, institutions, and systems.

METHODS

The research team included a chair, project manager, research fellow, and five research associates. The working group (N=19 members) included service users (N=3), and its members were selected to represent diverse professions and geographic areas. Ten areas of expertise were represented: psychometrists, psychiatrists, mental health nurses, psychologists, health economists, epidemiologists, national clinical quality programs, health service researchers, social workers, and service users, with members from 11 countries (Australia, Canada, Denmark, Greece, India, Israel, Japan, Mexico, the Netherlands, the United Kingdom, and the United States). No institutional review board approval or informed consent was necessary for this study.

Service User Focus Groups

Before commencement of the working group meetings, two focus groups were held with four service users to identify outcomes of highest importance to them. Service user recruitment occurred through recommendations from patient organizations and from working group members.

Systematic Literature Review to Identify Outcomes

Between January and March 2019, systematic literature searches were conducted to identify outcomes related to schizophrenia and bipolar disorder type I in adolescent and adult populations. The databases MEDLINE, PubMed, and PsycInfo were searched for publications between January 2013 and January 2019 (the search strategies are detailed in an online supplement to this article).

The Cochrane highly sensitive search strategy for identifying randomized trials (36) was first used to identify outcomes for schizophrenia spectrum and bipolar I disorders. Additional searches that excluded randomized trials were conducted in PsycInfo and MEDLINE. This included qualitative research that examined service users' perspectives on outcomes of importance and the impact of schizophrenia or bipolar disorder on service users' lives. Supplementary sources from working group members and reference lists in identified papers were used to find additional PROs and PROMs.

Consensus Process

A modified Delphi consensus process was used to select the outcome set (37). The process involved reaching consensus

in five main areas: scope, to determine which psychotic disorders, patient populations, and treatments to include; outcomes, to identify a minimum number of outcomes for inclusion in the set; measures, to assess each outcome; definitions and time points, to assess outcomes; and risk adjustment factors, to enable comparisons among providers implementing the set. The research team prepared and distributed presentations for review before each videoconferencing meeting. The five main areas were discussed during the meetings, and feedback from the working group was incorporated. Members rated and provided feedback on each item under review in the five main areas in online surveys.

On the basis of ICHOM processes outlined a priori, inclusion in the set required that at least 80% of the working group voted an item as "essential" (score of 7–9) in the first or second Delphi round (the scores in the Qualtrics survey ranged from 1, not recommended, to 9, essential). When consensus was not reached by voting, the item was discussed and revisited in the next meeting and survey. Outcomes were excluded if at least 80% of the working group voted an item as "not recommended" (score 1–3). The working group voted on all inconclusive outcomes in the third and final survey round, in which response options were "include" or "exclude" and inclusion required only 70% consensus.

Identification of Potential Outcome Measures From the Systematic Literature Review

Publications identified in the systematic literature review were the source of potential outcome measures. After development of a comprehensive measures list, a search filter was used to facilitate measure selection (38) that identified studies in PubMed reporting psychometric properties of measures. The research team excluded measures that had a cost associated with use.

Assessment of PROMs

The COnsensus-based Standards for the selection of health outcomes Measurement Instruments (COSMIN) checklist was used to assess the psychometric properties of measures, including reliability (test-retest and internal consistency), validity (content validity, face validity, and construct validity), and responsiveness (sensitivity to change) (14). In addition, the working group considered the feasibility of collecting the measures, including length of time to administer, training needed, and international applicability.

Breakout Sessions to Narrow the List of Measures

After the COSMIN checklist was applied, four breakout sessions were held to review and reduce the number of potential measures. The sessions were held with a small number of working group members with lived experience or professional expertise in the areas of functioning, personal recovery, symptoms, and treatment. Participants narrowed the list of potential measures and established a proposal for the wider working group to discuss and endorse. Any measures that took >20 minutes to complete were removed.

Risk Adjustment Factors

A preliminary list of risk adjustment factors and definitions was developed on the basis of risk adjustment factors identified from the systematic literature searches, national registries, and review of existing ICHOM standard sets. Factors were identified according to evidence of their effect on the outcomes selected. Demographic and clinical factors were assessed on relevance and feasibility of measurement. Further, to harmonize across ICHOM mental health standard sets, ICHOM mental health working group chairs developed a list of factors to reduce burden of implementation for service users with more than one diagnosis. The harmonized list was presented to each mental health consensus working group. Factors voted for inclusion reached 70% consensus.

Open Review and Patient Validation

After development, the set was distributed to international stakeholder groups, including professionals, adult service users, and caregivers in two separate stakeholder surveys to obtain feedback on outcomes, measures, risk adjustment factors, and timing of outcome collection. Respondents were recruited via networks identified by the research team and working group members through e-mail and social media and national and international patient organizations. Service users and caregivers were asked to rate the importance of each outcome using a 9-point Likert scale and were provided space to suggest additional outcomes.

RESULTS

Scope

The working group selected both schizophrenia spectrum disorders (as classified in *ICD-11* and *DSM-5*) and bipolar disorder type I (as classified in *ICD-11*). The set is limited to the adolescent (ages ≥12 years) and adult populations.

Service User Focus Groups

The core outcomes identified as important to service users included improvement in positive and negative symptoms, general medical health, and medication side effects; personal recovery; and prevention of relapse.

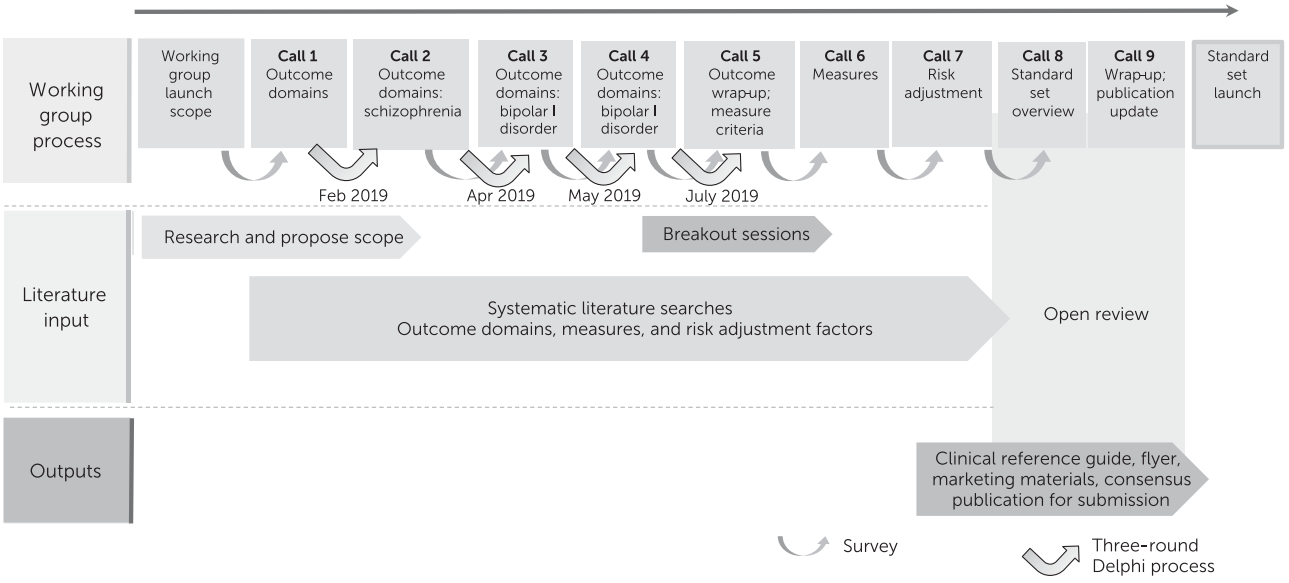
Identifying PROs

The literature searches identified 9,118 references, of which 758 were eligible for full-text review (see diagram in online supplement). Additional sources were recommended by working group members. After removal of duplicates and measures with associated costs and in languages other than English, a total of 131 measures were examined that assess symptoms, personal recovery, functioning, and treatment in psychosis (see online supplement).

Domain, Outcome, and Measure Selection by Consensus Working Group

The development process used by the consensus working group is illustrated in Figure 1. The working group identified

FIGURE 1. Summary of the development process for a standard set of patient-reported outcome measures for psychotic disorders



four outcome domains (symptoms, recovery, functioning, and treatment) that encompassed 14 outcomes (Table 1) (39–47). Symptoms included depressive symptoms, suicidal ideation and behavior, positive symptoms, negative symptoms (schizophrenia), mania or hypomania (bipolar I), sleep quality, and relapse rate as measured by hospitalizations. Personal recovery included quality-of-life measures. Functioning included global, social, and role functioning and general medical health. Medication side effects were included under treatment outcomes. Each core outcome identified by service users in the focus groups was included in the final set.

A total of 57 measures were presented to the working group for a vote (see online supplement). Outcome measures were reviewed in their entirety, with consideration given to psychometric properties, previous use in the specified population, the number and wording of items, and time taken to complete (Table 2) (39–50). Response rates were 80%, 85%, 80%, and 80% for the first through fourth modified Delphi processes, respectively.

Service users provided feedback regarding item wording and appropriateness and their opinion about the ability of the measure to capture the outcome for individuals with a psychotic disorder. This feedback was summarized in tables and presented alongside feedback from other working group members. A total of 147 measures were initially mapped to the 14 outcomes, and 39 measures were presented via short lists in breakout sessions.

Evaluation of Measures

Measures were assessed on psychometric properties, including acceptable interrater reliability ($r \geq 0.70$) (51), internal consistency (Cronbach's $\alpha \geq 0.70$) (52), and evidence of sensitivity to change, reflected by change in scores measured over time, consistent with a priori hypotheses about anticipated treatment outcome (Table 2). Measures rated as strong had acceptable interrater reliability, internal consistency, and evidence of sensitivity to change. Measures rated as mixed had only a single evaluation identified, mixed evidence from several sources, or strong evidence only for certain items. Measures rated as weak had below-threshold evidence. The working group selected nine measures in an outcome set that takes approximately 35 minutes to complete. The selected measures are freely available in English to users.

Time Point Recommendations

The outcome assessment time line was proposed by the working group to best achieve a balance between the clinically relevant times when outcomes may be expected to change and pragmatic concerns in data collection (53). The working group recommended assessment of outcomes before treatment as a baseline and then every 6 months and assessment of risk adjustment factors at baseline and annually thereafter.

Risk Adjustment Factors

The preliminary list of risk adjustment factors included demographic factors, such as age and socioeconomic status; clinical factors, such as comorbid conditions (54) and hospitalizations; and intervention factors, such as the setting. Harmonization of risk adjustment factors across mental health sets resulted in the addition of two factors: trauma, as assessed by adverse childhood experiences, and contact with law enforcement (Table 3).

Open Review and Patient Validation

Ninety-five professionals living in Australia, Canada, Chile, Nigeria, Sweden, the United Kingdom, and the United States responded to the open review survey. Service users and caregivers (N=25) were from Australia, the United Kingdom, and

TABLE 1. Domain, outcome, definition, measure, timing of administration, and data sources identified by the working group in developing a standard set of patient-reported outcome measures for psychotic disorders

Domain and outcome	Definition	Outcome measure[a]	Administration timing	Data source
Symptoms				
Depressive symptoms	Mood or emotional state that is marked by feelings such as depressed mood, hopelessness, worthlessness, or guilt and a reduced ability to enjoy life	PHQ-9 (41)	Baseline and every 6 months	Patient
Suicidal ideation and behavior	Suicidal ideation, suicidal thoughts or behaviors, suicide attempts, most often accompanied by intense feelings of hopelessness or depression or by self-destructive behaviors	PHQ-9 (41)	Baseline and every 6 months	Patient
Positive symptoms	Change in thoughts or perceptions, including hallucinations, delusions, or disorganized thought	MCSI (40)	Baseline and every 6 months	Patient
Negative symptoms[b]	A withdrawal or lack of function not expected in a healthy person, including blunting of affect, poverty of speech and thought, apathy, anhedonia, reduced social drive, loss of motivation, lack of social interest, and inattention to social or cognitive input	ReQoL-20 (39)	Baseline and every 6 months	Patient
Mania, hypomania[c]	Abnormally elevated mood state characterized by symptoms such as inappropriate elation, increased irritability, severe insomnia, grandiose notions, increased speed or volume of speech, disconnected and racing thoughts, increased energy and activity level, and inappropriate social behavior	ASRM (42)	Baseline and every 6 months	Patient
Sleep quality	Sleep problems resulting in decreased quality, including difficulty falling asleep, difficulty staying asleep, early morning awakening, and not feeling rested on waking up	PROMIS-Sleep (43)	Baseline and every 6 months	Patient
Relapse rate	Reemergence of symptoms or disorder after partial or complete recovery	Hospitalizations	Baseline and every 6 months	Clinician or patient
Recovery				
Personal recovery	A very personal process of changing one's attitudes, values, feelings, goals, skills, or roles. A way of living a satisfying, hopeful, and contributing life, according to the CHIME domains: connectedness, hope and optimism, identity, meaning and purpose, and empowerment	ReQoL-20 (39)	Baseline and every 6 months	Patient
Quality of life	Individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, including independence, experiencing a fuller range of emotions, and life satisfaction	ReQoL-20 (39)	Baseline and every 6 months	Patient
Functioning				
Global functioning	Individuals' social, occupational, and psychological functioning	WHODAS 2.0 (adult) (44); KIDSCREEN-10 (adolescent) (45)	Baseline and every 6 months	Patient
Social functioning	Individuals' interactions with their environment, the quality of those interactions, and their ability to fulfill their role within such environments as work, social activities, and relationships with partners, families, or friends	WHODAS 2.0 (adult) (44); KIDSCREEN-10 (adolescent) (45)	Baseline and every 6 months	Patient

continued

TABLE 1, *continued*

Domain and outcome	Definition	Outcome measure[a]	Administration timing	Data source
Role functioning	Ability to perform occupational activities or performance of functional tasks that support participation in the academic aspects of school	WHODAS 2.0 (adult) (44); KIDSCREEN-10 (adolescent) (45)	Baseline and every 6 months	Patient
Physical health	Measure of general medical health and well-being and overall satisfaction with general medical health	PHQ-15 (46)	Baseline and every 6 months	Patient
Treatment side effects	Effects of the prescribed medication, whether therapeutic or adverse, besides the intended treatment effect	GASS (47)	Baseline and every 6 months	Patient

[a] ASRM, Altman Self-Rating Mania Scale; GASS, Glasgow Antipsychotic Side-Effect Scale; MCSI, modified Colorado Symptom Index; PHQ-9, 9-item Patient Health Questionnaire; PHQ-15, 15-item Patient Health Questionnaire; PROMIS-Sleep, PROMIS Short Form V1.0 Sleep Disturbance 4a; ReQoL-20, 20-item version of Recovering Quality of Life; WHODAS 2.0, WHO Disability Assessment Schedule 2.0.
[b] Specific to schizophrenia spectrum disorders.
[c] Specific to bipolar I disorder.

TABLE 2. Psychometric properties of nine measures identified for the set of patient-reported outcome measures for psychotic disorders

Measure, abbreviation, and relevant study	No of items	Validity[a]	Reliability[a]	Sensitivity to change[b]
9-item Patient Health Questionnaire (PHQ-9) (41, 48, 49)	9	++	+	**
Modified Colorado Symptom Index (MCSI) (40)	14	+	++	×
20-item version of Recovering Quality of Life (ReQoL-20) (39, 50)	20	NA[c]	++	*
Altman Self-Rating Mania Scale (ASRM) (42)	5	NA	++	×
PROMIS Short Form V1.0 Sleep Disturbance 4a (PROMIS-Sleep) (43)	4	++	+	×
12-item version of the WHO Disability Assessment Schedule 2.0 (WHODAS 2.0 [adult]) (44)	12	+	+	×
10-item KIDSCREEN (adolescent) (45)	10	++	+	×
15-item Patient Health Questionnaire (PHQ-15) (46, 49)	15	+	++	*
Glasgow Antipsychotic Side-Effect Scale (GASS) (47)	22	++	++	×

[a] ++, strong: acceptable interrater reliability ($r \geq 0.70$) and internal consistency (Cronbach's $\alpha \geq 0.70$) across most studies identified; +, mixed: only a single evaluation identified, mixed evidence from several sources, or strong evidence only for certain items or sections; NA, not assessed.
[b] **, well-established sensitivity to change; *, emerging evidence of sensitivity to change; ×, further studies needed to assess sensitivity to change.
[c] Not a validated measure of negative symptoms; includes items that assess negative symptoms.

the United States. Overall endorsement of the set and its elements exceeded the required 70%. Of the 106 participants with professional experience, support for outcome domains ranged from 77% to 93%, and support for outcome measures ranged from 61% to 84%. Of the 25 participants with lived experience, 92% agreed that the measures are useful to collect, and 91% stated that the list captured all important outcomes. Endorsement of included outcomes ranged from 72% to 92%.

DISCUSSION

The research team identified numerous PROMs for both schizophrenia spectrum disorders and bipolar disorder. The process involved a review to identify the measures and a subsequent review to assess the measures' psychometric properties (55), a process similar to that used by other reviewers of PROMs. The consensus process was successful in reducing the number of measures to a pragmatic set for use in routine practice. Service users were integral to the development of the set, from initial identification of core outcomes that mattered most to them to assessment of measures' face validity and comprehensibility of items. The open review and patient validation phase helped ensure the interpretability and cultural sensitivity of the set.

In addition to the well-recognized symptoms associated with psychosis, sleep problems are common among people with schizophrenia spectrum disorders and bipolar I disorder; they have a negative impact on functioning and well-being and are associated with a reduced ability or opportunity to participate in valued activities (56). Sleep quality, as assessed by the PROMIS-Sleep measure (43), is included in the set.

Personal recovery was very important to service users. The personal recovery measure in the outcome set has good psychometric properties and has been used in published research on populations of mental health service users (39, 57, 58). We did not find a positive symptom measure that has been developed and tested exclusively in samples of people with

schizophrenia. However, the modified Colorado Symptom Index has been used in large-scale studies of populations with severe mental illness (40). We did not identify a self-report measure for negative symptoms. Because of its length and the availability of only one language version, the Clinical Assessment Interview for Negative Symptoms (59), a 30-item measure, was not recommended for inclusion in the outcome set. As a best alternative, the working group suggested a PROM, the Recovering Quality of Life (39). Historically, a clinician-rated outcome measure (CROM), the Quality of Life Scale (60), has been used as a negative symptom measure. There was less research to support decision making on adolescent measures.

Limitations of this work include a low number of service users in the working group and no service users with lived experience of bipolar I disorder. The inclusion of service users in developing PROMs is important yet remains challenging. Few studies include them at all stages of development (61). In this study, service users were recruited after the design stage and before the decision to include PROMs for bipolar disorder. At project commencement, two service user–only focus groups were held. Additionally, we paid specific attention to recovery measure studies that involved service users in their development and evaluation. A limitation of the final set is redundancy in some items across measures. For example, sleep is assessed in the PROMIS-Sleep measure and in the depression and mania measures. This commonly encountered issue could be

TABLE 3. Risk adjustment factors identified by the working group in developing a standard set of patient-reported outcome measures for psychotic disorders

Risk adjustment area, patient population, and measure	Supporting information[a]	Administration timing	Data source
Demographic factor			
All patients			
Year of birth	NA	Baseline	Patient reported
Sex	Sex at birth	Baseline	Patient reported
Gender identity	NA	Baseline	Patient reported
Sexual orientation	NA	Baseline	Patient reported
Socioeconomic status	Adults, highest level of education completed; adolescents, highest level of education completed by parents (proxy to be used)	Baseline, transition to adult services, and annually if still in education	Patient reported
Work or education status	NA	Baseline and annually	Patient reported
Housing status	NA	Baseline and annually	Patient reported
Living arrangement	NA	Baseline and annually	Patient reported
Ethnic minority group or marginalization	NA	Baseline	Patient reported
Adult patients and adolescent patients where appropriate			
Contact with law enforcement	To be administered to adolescents only when appropriate to do so and for whom this measure would not cause unnecessary distress. Baseline, ever been convicted (lifetime); annually, ever been convicted (in past 12 months)	Baseline and annually	Patient reported
Clinical factor			
All patients			
Comorbid conditions	Based on the Self-Administered Comorbidity Questionnaire (54)	Baseline and annually	Patient reported
Hospitalizations	Number of lifetime hospitalizations related to the target condition	Baseline	Administrative data
Adult patients and adolescent patients where appropriate			
Adverse life experiences	To be administered to adolescents only when appropriate to do so and for whom this measure would not cause unnecessary distress	Baseline and transition to adult services	Patient reported
Intervention factor			
All patients			
Intervention setting	NA	Baseline and annually	Clinical
Intervention type	NA	Baseline and annually	Clinical

[a] NA, not applicable.

addressed in future research by using statistical methods to address overlap across the entire outcome set. Consistent with a systematic review of PROMs and CROMs for assessing youth outcomes (62), we found broader outcome measures developed for adolescents, including measures for quality of life. Targeted measures for symptoms and treatment side effects were often developed for adults and rarely tested with or adapted for adolescents. This research gap highlights the need to validate outcome measures for the adolescent population.

Important psychometric properties were not included in our selection criteria. These properties included meaningful change thresholds (63) and severity thresholds, such as mild, moderate, and severe, which can be linked to treatment decisions (64). The properties were present to varying degrees in the selected measures, but they were not used as selection criteria. A critical evaluation of each measure was therefore beyond the scope of this project. We did not address alternative modes of PRO administration, an important consideration in implementation. However, a meta-analytic review concluded substantial evidence indicating the equivalence of computer- and paper-administered PROs (65).

CONCLUSIONS

A standardized set of nine PROMs was identified in this study. The set can be used to support measurement-based care and, in combination with risk adjustment factors, to compare program outcomes. Finally, the set can be used to support development of value-based health care for people with psychotic disorders.

AUTHOR AND ARTICLE INFORMATION

Department of Psychiatry, University of Calgary, Calgary, Canada (McKenzie, Addington); Department of Zoology, University of Oxford, Oxford, United Kingdom (Matkin); International Consortium for Health Outcomes Measurement, London (Sousa Fialho, Emelurumonye, Gintner, Ilesanmi, Jagger, Quinney); Service user, Calgary, Canada (Anderson); Mental Health Centre, Copenhagen (Baandrup); Service user, Pune, India (Bakhshy); Tees, Esk and Wear Valleys National Health Service (NHS) Foundation Trust, Darlington, United Kingdom (Brabban); Australian Mental Health Outcomes and Classification Network, St. Leonards, Australia (Coombs); Department of Psychiatry and Molecular Medicine, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York, and Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin (Correll); South London and Maudsley NHS Foundation Trust, London (Cupitt); School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom (Keetharuth); Department of Biomedical Informatics, Universidad Nacional Autónoma de México, Mexico City (Lima); Institute for Lifecourse Development, University of Greenwich, and King's Health Economics, King's College, London (McCrone); School of Nursing, Pacific Lutheran University, Tacoma, Washington (Moller); Department of Psychiatry, Erasmus Medical Center, University Medical Center, Rotterdam, Netherlands (Mulder); Department of Community Mental Health, University of Haifa, Haifa, Israel (Roe); Northern Clinical School, Faculty of Medicine and Health, University of Sydney, Sydney, and System Information and Analytics Branch, New South Wales Ministry of Health, St. Leonards, Australia (Sara); Cochrane Schizophrenia Group, London (Shokraneh); City University of London, London (Sin); Center for Psychiatric Research, Maine Medical Center Research Institute, Scarborough, and Department of Psychiatry, Tufts School of Medicine, Boston (Woodberry). Send correspondence to Ms. McKenzie (emckenzi@ucalgary.ca).

Dr. Correll reports services as a consultant or advisor to or receipt of honoraria from Acadia, Alkermes, Allergan, Angelini, Axsome, Gedeon Richter, Gerson Lehrman Group, Indivior, IntraCellular Therapies, Janssen/J&J, Karuna, LB Pharma, Lundbeck, MedAvante-ProPhase, MedInCell, Medscape, Merck, Mylan, Neurocrine, Noven, Otsuka, Pfizer, Recordati, Rovi, Servier, Sumitomo Dainippon, Sunovion, Supernus, Takeda, and Teva. He also reports provision of expert testimony for Janssen and Otsuka; service on data safety monitoring boards for Lundbeck, Rovi, Supernus, and Teva; receipt of grant support from Janssen and Takeda; and being a stock option holder of LB Pharma.

Received December 5, 2020; revisions received March 25 and May 7, 2021; accepted May 13, 2021; published online August 9, 2021.

REFERENCES

1. US Department of Health and Human Services: Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006; 4:79
2. Greenhalgh J: The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res 2009; 18:115–123
3. Sisodia RC, Dankers C, Orav J, et al: Factors associated with increased collection of patient-reported outcomes within a large health care system. JAMA Netw Open 2020; 3:e202764
4. Acquadro C, Berzon R, Dubois D, et al: Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. Value Health 2003; 6:522–531
5. Evans JP, Smith A, Gibbons C, et al: The National Institutes of Health Patient-Reported Outcomes Measurement Information System (PROMIS): a view from the UK. Patient Relat Outcome Meas 2018; 9:345–352
6. Cella D, Riley W, Stone A, et al: The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010; 63:1179–1194
7. Lavallee DC, Chenok KE, Love RM, et al: Incorporating patient-reported outcomes into health care to engage patients and enhance care. Health Aff 2016; 35:575–582
8. Eremenco S, Pease S, Mann S, et al: Patient-Reported Outcome (PRO) Consortium translation process: consensus development of updated best practices. J Patient Rep Outcomes 2017; 2:12
9. Clark DM, Canvin L, Green J, et al: Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. Lancet 2018; 391:679–686
10. Burgess P, Pirkis J, Coombs T: Routine outcome measurement in Australia. Int Rev Psychiatry 2015; 27:264–275
11. Kendrick T, El-Gohary M, Stuart B, et al: Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. Cochrane Database Syst Rev 2016; 7:CD011119
12. Turner RR, Quittner AL, Parasuraman BM, et al: Patient-reported outcomes: instrument development and selection issues. Value Health 2007; 10(suppl 2):S86–S93
13. Valderas JM, Ferrer M, Mendívil J, et al: Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. Value Health 2008; 11:700–708
14. Mokkink LB, Prinsen CA, Bouter LM, et al: The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. Braz J Phys Ther 2016; 20:105–113

15. Snyder CF, Aaronson NK, Choucair AK, et al: Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Qual Life Res 2012; 21: 1305–1314
16. Vigo D, Thornicroft G, Atun R: Estimating the true global burden of mental illness. Lancet Psychiatry 2016; 3:171–178
17. Saha S, Chant D, McGrath J: Meta-analyses of the incidence and prevalence of schizophrenia: conceptual and methodological issues. Int J Methods Psychiatr Res 2008; 17:55–61
18. Charlson FJ, Ferrari AJ, Santomauro DF, et al: Global epidemiology and burden of schizophrenia: findings from the Global Burden of Disease Study 2016. Schizophr Bull 2018; 44:1195–1203
19. Hjorthøj C, Stürup AE, McGrath JJ, et al: Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. Lancet Psychiatry 2017; 4:295–301
20. Jongsma HE, Turner C, Kirkbride JB, et al: International incidence of psychotic disorders, 2002–17: a systematic review and meta-analysis. Lancet Public Health 2019; 4:e229–e244
21. Firth J, Siddiqi N, Koyanagi A, et al: The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. Lancet Psychiatry 2019; 6:675–712
22. Correll CU, Solmi M, Veronese N, et al: Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. World Psychiatry 2017; 16:163–180
23. Vancampfort D, Stubbs B, Mitchell AJ, et al: Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. World Psychiatry 2015; 14:339–347
24. Ferrari AJ, Stockings E, Khoo JP, et al: The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. Bipolar Disord 2016; 18:440–450
25. Moreira ALR, Van Meter A, Genzlinger J, et al: Review and meta-analysis of epidemiologic studies of adult bipolar disorder. J Clin Psychiatry 2017; 78:e1259–e1269
26. Jääskeläinen E, Juola P, Hirvonen N, et al: A systematic review and meta-analysis of recovery in schizophrenia. Schizophr Bull 2013; 39:1296–1306
27. Leamy M, Bird V, Le Boutillier C, et al: Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011; 199:445–452
28. Correll CU, Galling B, Pawar A, et al: Comparison of early intervention services vs treatment as usual for early-phase psychosis: a systematic review, meta-analysis, and meta-regression. JAMA Psychiatry 2018; 75:555–565
29. Morrison AP, Law H, Barrowclough C, et al: Psychological approaches to understanding and promoting recovery in psychosis and bipolar disorder: a mixed-methods approach. Programme Grants Appl Res 2016; 4; doi: 10.3310/pgfar04050
30. Meadows G, Brophy L, Shawyer F, et al: REFOCUS-PULSAR recovery-oriented practice training in specialist mental health care: a stepped-wedge cluster randomised controlled trial. Lancet Psychiatry 2019; 6:103–114
31. Addington D, Anderson E, Kelly M, et al: Canadian practice guidelines for comprehensive community treatment for schizophrenia and schizophrenia spectrum disorders. Can J Psychiatry 2017; 62:662–672
32. Namjoshi MA, Buesching DP: A review of the health-related quality of life literature in bipolar disorder. Qual Life Res 2001; 10:105–115
33. Chen M, Fitzgerald HM, Madera JJ, et al: Functional outcome assessment in bipolar disorder: a systematic literature review. Bipolar Disord 2019; 21:194–214
34. Porter ME, Larsson S, Lee TH: Standardizing patient outcomes measurement. N Engl J Med 2016; 374:504–506
35. Obbarius A, van Maasakkers L, Baer L, et al: Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM Depression and Anxiety Working Group. Qual Life Res 2017; 26:3211–3225
36. Cumpston M, Li T, Page MJ, et al: Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019; 10:ED000142
37. Fitch K, Bochner A, Keller DS: Cost comparison of laparoscopic colectomy versus open colectomy in colon cancer. Curr Med Res Opin 2017; 33:1215–1221
38. Terwee CB, Jansma EP, Riphagen II, et al: Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009; 18:1115–1123
39. Keetharuth AD, Brazier J, Connell J, et al: Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. Br J Psychiatry 2018; 212:42–49
40. Conrad KJ, Yagelka JR, Matters MD, et al: Reliability and validity of a modified Colorado Symptom Index in a national homeless sample. Ment Health Serv Res 2001; 3:141–153
41. Kroenke K, Spitzer RL, Williams JB: The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001; 16: 606–613
42. Altman EG, Hedeker D, Peterson JL, et al: The Altman Self-Rating Mania Scale. Biol Psychiatry 1997; 42:948–955
43. Buysse DJ, Yu L, Moul DE, et al: Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. Sleep 2010; 33:781–792
44. Ustun TBKN, Chatterji S: Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0). Geneva, World Health Organization, 2010
45. Ravens-Sieberer U, Erhart M, Rajmil L, et al: Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. Qual Life Res 2010; 19:1487–1500
46. Kroenke K, Spitzer RL, Williams JB: The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002; 64:258–266
47. Waddell L, Taylor M: A new self-rating scale for detecting atypical or second-generation antipsychotic side effects. J Psychopharmacol 2008; 22:238–243
48. Beard C, Hsu KJ, Rifkin LS, et al: Validation of the PHQ-9 in a psychiatric sample. J Affect Disord 2016; 193:267–273
49. Kroenke K, Spitzer RL, Williams JB, et al: The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010; 32:345–359
50. Keetharuth A, Brazier J, Connell J, et al: Development and Validation of the Recovering Quality of Life (ReQoL) Outcome Measures. EEPRU Research Report 050. Sheffield, United Kingdom, Policy Research Unit in Economic Evaluation of Health and Care Interventions, Universities of Sheffield and York, 2017
51. Reeve BB, Wyrwich KW, Wu AW, et al: ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res 2013; 22:1889–1905
52. Nunnally JC: Psychometric Theory. New York, McGraw-Hill, 1978
53. Zerillo JA, Schouwenburg MG, van Bommel ACM, et al: An international collaborative standardizing a comprehensive patient-centered outcomes measurement set for colorectal cancer. JAMA Oncol 2017; 3:686–694
54. Sangha O, Stucki G, Liang MH, et al: The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003; 49:156–163

55. Prinsen CAC, Mokkink LB, Bouter LM, et al: COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018; 27:1147–1157
56. Faulkner S, Bee P: Experiences, perspectives and priorities of people with schizophrenia spectrum disorders regarding sleep disturbance and its treatment: a qualitative study. BMC Psychiatry 2017; 17:158
57. Connell J, Carlton J, Grundy A, et al: The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). Qual Life Res 2018; 27:1893–1902
58. Keetharuth AD, Taylor Buck E, Acquadro C, et al: Integrating qualitative and quantitative data in the development of outcome measures: the case of the Recovering Quality of Life (ReQoL) measures in mental health populations. Int J Environ Res Public Health 2018; 15:15
59. Park SG, Llerena K, McCarthy JM, et al: Screening for negative symptoms: preliminary results from the self-report version of the Clinical Assessment Interview for Negative Symptoms. Schizophr Res 2012; 135:139–143
60. Mueser KT, Kim M, Addington J, et al: Confirmatory factor analysis of the Quality of Life Scale and new proposed factor structure for the Quality of Life Scale–Revised. Schizophr Res 2017; 181:117–123
61. Wiering B, de Boer D, Delnoij D: Patient involvement in the development of patient-reported outcome measures: a scoping review. Health Expect 2017; 20:11–23
62. Kwan B, Rickwood DJ: A systematic review of mental health outcome measures for young people aged 12 to 25 years. BMC Psychiatry 2015; 15:279
63. Jacobson NS, Truax P: Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991; 59:12–19
64. Jaeschke R, Singer J, Guyatt GH: Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989; 10:407–415
65. Gwaltney CJ, Shields AL, Shiffman S: Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value Health 2008; 11: 322–333

Submissions Invited for Social Determinants of Mental Health Column

Coeditors: Ruth S. Shim, M.D., M.P.H., and Michael T. Compton, M.D., M.P.H.

This column focuses on clinical and policy issues as they relate to social justice in psychiatry and the social determinants of mental health, with a specific focus on mental health disparities and evidence-based strategies to improve mental health equity across population groups. Initiatives taking place in hospitals, clinics, health systems, and insurance plans are emphasized. Ways in which clinicians and mental health services can address (screen for, evaluate, and ameliorate) social determinants of mental health are highlighted. Manuscripts that emphasize specific social determinants of mental health, including discrimination, adverse early life experiences, poverty, social exclusion, low employment status, and low educational attainment, to name a few—and particularly how these determinants connect to mental health outcomes and can be addressed by mental health services—are particularly welcome. Submissions (via mc.manuscriptcentral.com/appi-ps) are limited to 2,400 total words, inclusive of a 100-word abstract, two or three one-sentence Highlights, and up to 10 references.