Measuring patient-reported outcomes in psychosis: conceptual and methodological review

Ulrich Reininghaus and Stefan Priebe

Background
There are calls to use patient-reported outcomes (PROs) routinely across mental health services. However, the use of PROs in patients with psychosis has been questioned.

Aims
To examine the concepts and measures of four widely used PROs: treatment satisfaction, subjective quality of life, needs for care and the quality of the therapeutic relationship.

Method
We conducted a literature search of academic databases on concepts, characteristics and psychometric properties of the four PROs in patients with psychosis.

Results
Although numerous concepts and measures have been published, evidence on the methodological quality of existing PROs is limited. Measures designed to assess distinct PROs showed a considerable conceptual, operational and empirical overlap, and some of them also included specific aspects. The impact of symptoms and cognitive deficits appears unlikely to be of clinical significance.

Conclusions
The popularity of PROs has not been matched with progress in their conceptualisation and measurement. Based on current evidence, some recommendations can be made. Distinct and short measures with clinical relevance and sufficient psychometric properties should be preferred. Future research should optimise the validity and measurement precision of PROs, while reducing assessment burden.

Declaration of interest
None.

Method
A review of the conceptual and methodological literature on the four PROs in the care of patients with psychosis was conducted. We searched the literature systematically and also followed the recommendations for conceptual and methodological reviews to search widely in disparate sources and allow for overlap in the various stages (literature search, analysis and writing).

Search strategy and selection criteria
A search of the academic databases EMBASE, Medline and PsycINFO was performed to identify papers that, first, reported the characteristics and psychometric properties of PRO measures to assess treatment satisfaction, SQoL, needs for care and the therapeutic relationship in the care of patients with psychosis, and second, provided definitions of concepts intended to be assessed by at least one of the identified measures. The term ‘PRO’ was used in accordance with the FDA definition given earlier.

Search criteria
We searched the literature systematically and also followed the recommendations for conceptual and methodological reviews to search widely in disparate sources and allow for overlap in the various stages (literature search, analysis and writing).

Selection criteria
(p. 2).

Since the 1960s, patient-reported outcomes (PROs) have become increasingly popular in the care of patients with psychosis. There is no universally accepted terminology and definition of such outcomes. In the literature the terms ‘PROs’, ‘patient-reported outcome measures’ (PROMs), ‘patient-based outcomes’, ‘patient-driven outcomes’, ‘self-rated outcomes’ and ‘subjective evaluation criteria’ have been used interchangeably. In recent years the term ‘PRO’ appears to be most widely used. The US Food and Drug Administration (FDA) defined PROs as:

‘any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else’ (p. 2).

Treatment satisfaction, subjective quality of life (SQoL), needs and the quality of the therapeutic relationship can be considered as four historically rooted, commonly used and important PRO concepts in the care of patients with psychosis. Whereas the list of PROs has increased steadily, their popularity has gained momentum over the past decade, partly through their intuitive appeal for stakeholder groups. In the UK a recent National Health Service (NHS) White Paper announced plans for new outcome assessments, in which PROs were to be used to measure the effectiveness of services. Using PROs in the monitoring of outcomes of individual patients and services can also feed into the patient–clinician communication, reflective practice, quality management and service development. However, the selection of appropriate concepts and measures often remains difficult. Further, some authors have questioned the use of PROs in patients with psychosis owing to conceptual and methodological shortcomings, with some proposing to discard them entirely. Against this background, this review aimed to examine the concepts and measures of four widely used PROs – treatment satisfaction, SQoL, needs for care and the quality of the therapeutic relationship – in the evaluation of care of patients with psychosis.
of academic databases, informal networks were used to identify papers.

Data extraction and synthesis

As the conceptual and methodological literature on PROs in the evaluation of treatments for psychosis is vast and disparate, a quantitative synthesis appeared neither appropriate nor feasible. The findings are presented descriptively. Although PRO measures can be distinguished according to various characteristics, we focused on the following ones: concept purported to be measured, number and content of domains, estimated completion time, response options and type (generic, condition- or disease-specific, treatment-specific and utility measures). Various psychometric properties for evaluating PROs have been proposed in the literature. We distinguished between reliability (i.e. internal consistency, reliability, scale information), validity (i.e. content validity, including face validity, and construct validity, including structural, convergent, discriminant, cross-cultural, concurrent and predictive validity) and responsiveness. Given the lack of consensus on how these psychometric properties are best evaluated and founds synthesised, we used a simple, dichotomous rating of whether or not a psychometric property had been examined for a given instrument.

Results

The results of the search strategy are summarised in Fig. 1. The search initially yielded a total of 2181 items (813 duplicates). Titles and abstracts were screened for 1368 references. Based on title and abstract sifts, 1238 references were excluded because they did not focus on the four PROs or psychosis. The number of potentially relevant references increased from 130 to 224 when additional items were added. Of these, 49 references were excluded for different reasons. Hence, from the 2181 initially identified references, only 175 were included in our review.

Concepts and definitions

Definitions of concepts to be assessed by the identified PRO measures were found in the literature. Definitions of treatment satisfaction, SQoL, needs for care and the therapeutic relationship are summarised in online Table DS1. Definitions of treatment satisfaction for which a definition of the concept to be measured was provided all purported to assess the multidimensional satisfaction concept of a personal evaluation of healthcare services and providers as proposed by Ware et al and Ruggeri et al. The identified SQoL measures were intended to assess a range of concepts (Table DS1). The only measure of needs that provided a definition of the concept to be measured, the Camberwell Assessment of Need (CAN), purported to assess a supply and perceived need concept. Pan-theoretical, Rogerian, systemic and psychoanalytic concepts were intended to be assessed by the identified measures of the therapeutic relationship. For each of the four PROs, no single universally accepted definition could be identified. Nevertheless, there were attempts to identify a common conceptual basis. Lauer noted that:

"There is agreement that quality of life is a multi-dimensional phenomenon and construct, aiming at a holistic or global perspective of individuals in their biopsychosocial nature" (p. 19).

Similarly, Ware et al emphasised that treatment satisfaction is most widely measured as a multidomain concept. However, this may imply a risk of providing non-specific or overinclusive definitions. Several PRO concepts other than the one to be measured may meet very broad definitions; for example, Stevens & Gabbay defined needs as ‘the ability to benefit in some way from health care’ (p. 21). Others have found a lack of clarity as to precisely the nature of some PRO concepts:

"In psychiatry, there is as yet no clearly defined concept of the therapeutic alliance" (Catty: p. 265).

A tendency was found to use terms from different theoretical backgrounds and traditions with at least slightly different connotations synonymously. For example, the term ‘therapeutic relationship’ has been used interchangeably with the terms ‘therapeutic alliance’, ‘helping alliance’ or ‘working alliance’, each of which has emerged from different lines of research. "Treatment satisfaction" has been used synonymously with ‘patient satisfaction’, ‘service satisfaction’ and ‘satisfaction with care’, to name a few. This may lead to a lack of clarity as to precisely which conceptualisation of PROs is being referred to. Several definitions of PRO concepts were found to overlap with others (Table DS1). However, some definitions of PRO concepts did not, and contained specific elements: this applied to definitions of SQoL, needs for care, and the therapeutic relationship. Overall, definitions of PRO concepts were found to vary in the extent to which they included overlapping and specific aspects.

Characteristics of PRO measures

Findings on characteristics and psychometric properties of PRO measures to assess treatment satisfaction, SQoL, needs for care and the therapeutic relationship are summarised respectively in online Tables DS2–DS5. For several measures the concept that the measure was intended to assess was not provided. Most
measures were generic in nature and used Likert scales. Short versions have been developed for several measures, based on conceptual and practical rather than empirical considerations. A number of measures were found to be long and time-consuming to administer: several had more than 30 items and a completion time greater than 20 min.

Several PRO measures were intended to assess multidomain concepts, with items being grouped within domains, and domains within more general PRO concepts. An overlap in the content of domains was observed across measures that were intended to assess different PROs. Specifically, the domains of measures to assess SQoL are similar and in part even identical to domains included in measures of needs. This applies likewise to measures of treatment satisfaction and the therapeutic relationship. The content of domains of treatment satisfaction and needs for care measures, and the content of treatment satisfaction and SQoL measures, show substantial overlap (Tables DS2–DS5).

Psychometric properties of PRO measures

The evaluation of the reviewed measures often included only limited information on psychometric properties in patients with psychosis (Tables DS2–DS5). The methods used to assess structural validity were largely not appropriate for ordinal data, as required for the predominantly used Likert scales. Only for two measures, the Quality of Life Interview (QoLI) and EuroQoL-5D, there was evidence on structural validity based on confirmatory factor analysis for ordinal data or item response modelling. For most measures there was no evidence on their measurement precision throughout the full range of scores; only for the QoLI was this psychometric property examined. For some measures no evidence on their internal consistency, test–retest reliability and scale information as well as content, structural, discriminant, convergent, concurrent, predictive or cross-cultural validity was found in the included studies.

Empirical overlap of PRO measures

Only a few studies assessed more than one outcome at a time. They consistently suggest low discriminant validity due to an empirical overlap of measures designed to assess different outcomes. The outcomes were substantially correlated and a single general factor explained more than half of the variance in SQoL, needs for care and treatment satisfaction scores. The general factor has been interpreted as a general appraisal tendency, but also shows the relevance of concept-specific aspects. The latter provide distinct information independent from both the general appraisal tendency and other concepts. Accurate measurement of changes in PROs provides a strong argument for the importance of cross-cultural validity in PRO measures, which was not systematically assessed in the reviewed studies.

Association with psychiatric symptoms and cognitive deficits

There was also evidence from several studies that less favourable SQoL is related to higher levels of psychopathologic disorder including positive, negative and depressive symptoms. For the association of psychiatric symptoms and treatment satisfaction, Katsakou & Priebe reported an inverse relationship between psychiatric symptoms and level of treatment satisfaction, which is in line with other studies. There are also a number of studies suggesting that patients with more severe psychotic symptoms have more unmet and total needs for care. However, a more recent pooled analysis of individual patient-level data obtained from 16 studies found that symptom levels were less strongly associated with SQoL in schizophrenia than in other mental disorders. A pooled analysis of associations between changes of symptoms and SQoL ratings over time identified an explained variance of only 5.5%. With respect to cognitive deficits, evidence on associations with PROs remains inconsistent. Fujii et al found that better cognitive performance was associated with lower SQoL ratings in a prospective study of patients with severe and enduring psychosis, which is consistent with other studies.

Discussion

Our review examining concepts and measures of four established PROs in the evaluation of treatments for psychosis generated at least three important findings. First, despite the increasing popularity of PROs with numerous concepts and measures, evidence of their methodological quality remains limited. Second, there is a considerable conceptual, operational and empirical overlap across measures designed to assess different PROs, although some concepts and measures also included aspects specific to individual PROs. Last, the influence of (or bias by) cognitive deficits and psychiatric symptoms appears limited and unlikely to be of clinical significance.

Limitations

The review has several limitations. The findings may be biased, as important references on concepts, characteristics and psychometric properties of PRO measures may have been missed. Concepts that might be relevant for one of the four PROs, but were not captured in an existing measure, were not included. The review was selective in examining concepts and measures of only four PROs and only a limited number of psychometric properties. Although Mokkink et al achieved a degree of consensus on the terminology and definitions of psychometric properties and provided guidance on data synthesis for reviews of the methodological quality of studies investigating psychometric properties of PROs, there is no consensus on how to synthesise findings on psychometric properties per se. We classified PROs according to whether or not they assessed specific psychometric properties. Given the absence of a consensus, this did not include ratings of the extent to which these psychometric properties were met. Finally, given the nature of conceptual and methodological reviews, there may have been a subjective bias of the authors in the analysis and interpretation of the literature.

Methodological quality of PROs

Over the past decades numerous concepts and measures of PROs have emerged. In contrast, our review found only limited evidence of their methodological quality. Several measures were not linked to specific concepts. A number of measures were long...
and time-consuming to administer. This may imply undue assessment burden on patients with psychosis as well as increased assessment costs. For most measures there was no evidence on their measurement precision throughout the full range of scores, as has been established by a few studies for observer-rated outcome measures in mental health.78 and, on a larger scale, for PROs in other medical disorders.79 The methods used to assess structural validity were largely not appropriate for ordinal data.80 Only a few of the reviewed studies conducted analyses based on confirmatory factor analysis for ordinal data or item response modelling.80,81 There are several implications of treating ordinal data as continuous, including attenuated relationships among PRO items in the presence of floor or ceiling effects, presence of pseudofactors and incorrect parameter estimates.82 These may challenge findings on the structural validity of PRO measures. In other words, measures using Likert scales, which have not been examined with psychometric methods appropriate for ordinal data, may be impaired in their ability to summarise patients’ item responses into scores that adequately reflect their dimensional structure. This is, however, central for the use of PROs in the evaluation of care, as such scores provide the basis on which value is assigned to treatments.

Conceptual, operational and empirical overlap

The conceptual, operational and empirical overlap of PROs has several implications for the validity of existing PRO measures. Campbell & Fiske, in their seminal work on discriminant and convergent validity, stated: ‘One cannot define without implying distinctions, and the verification of these distinctions is an important part of the validational process’ (p. 84).83

The verification of distinctions appears to be a part of the validational process that has been neglected by most of the research into PROs. New concepts were often proposed without assessing whether they were sufficiently distinct from existing concepts to warrant them being measured separately. This review suggests that an insufficient distinction between PROs at the conceptual level has led to a considerable overlap in the content of specific domains. This implies that, both at a conceptual and operational level, the requirements for establishing discriminant validity were not sufficiently considered when developing PROs. Empirically this may limit the ability of established measures to capture variance specific to the given concept. Indeed, this points towards substantial empirical overlap across measures. Although such overlap may reflect real associations between different PROs (e.g. one PRO influencing another), it still impairs the ability of each PRO measure to capture distinct information and, in psychometric terms, their discriminant validity.83

However, some concepts and operationalisations included aspects that were specific to one or more PROs. Recent evidence suggests that PROs may reflect both a general appraisal tendency that uniformly influences all PRO ratings in a positive or negative direction and components that are specific for each PRO. The specific information is independent of the general appraisal tendency. Maximising the specific information may be a challenge for future scale improvements.

Influence of cognitive deficits and psychiatric symptoms

In contrast to the concerns of some authors that the validity of existing PRO measures might be impaired owing to the influence of psychiatric symptoms and cognitive deficits,5,16 findings from our review suggested that the influence of (or bias by) cognitive deficits and psychiatric symptoms is very limited. The identified associations of PROs with symptoms and deficits do not compromise their validity as independent outcome criteria. However, all the evidence was taken from patients who consented to participate in research and were seen as capable of providing reasonable responses. Patients with high symptom levels may have been excluded from such studies, by clinicians or researchers. There is no evidence of a possible threshold of general or specific symptoms above which PROs might yield less reliable results.

Routine use of PROs

The conceptualisation and measurement of PROs in patients with psychosis are of practical relevance. These measures have an intuitive appeal for various stakeholder groups and there are calls to use them routinely across mental health services.8 Even though evidence on the methodological quality of PROs is limited overall, there are at least five recommendations that can be made about the routine use of PROs in the evaluation of treatments for psychosis.

(a) It should be carefully considered which PRO is relevant to the aim and approach of the given service, and what the implications of its results would be for service delivery and development.
(b) The use of several PRO measures should be avoided unless they address clearly distinct domains.
(c) Measures with evidence of good psychometric properties should be preferred; the evidence on psychometric properties is limited for most measures. Overall, measures using satisfaction-based concepts (e.g. assessing satisfaction with life domains or with treatment) have been more rigorously studied than others.
(d) In the absence of evidence showing that longer measures have superior properties, shorter measures should be prioritised to minimise the burden and costs of measurement. However, longer measures tend to be more reliable, and there can be a trade-off between brevity and psychometric quality.
(e) The influence of symptoms and cognitive deficits is unlikely to affect findings in small samples (although even a small explained variance may be relevant for research in large samples).

Future research

Despite the popularity of PROs for measuring the quality of routine mental healthcare, there are a number of conceptual and methodological shortcomings. Although according to our main findings this includes considerable conceptual, operational and empirical overlap across measures designed to assess different PROs, the influence of cognitive deficits and psychiatric symptoms appears limited. There is a need for more rigorous research to identify short measures that assess distinct PROs independent from overlap with highest possible precision. New methods such as item response modelling, item banking and computerised adaptive testing may help move this forward.74,79 Although such methods have been infrequently used in psychiatric studies,78 they have led to progress in measuring PROs in other medical conditions. A prominent example is the Patient-Reported Outcomes Measurement Information System (PROMIS).79 Computerised adaptive testing iteratively selects the item providing the highest precision for a given patient until a desired level of precision is achieved. This minimises the number of items each patient has to complete,84 and can be implemented on handheld electronic
devices. Ideally, conceptual and methodological work should be linked in future research to advance the measurement of PROs in patients with psychosis, so that concepts can both inform research and be refined on the basis of empirical data.

**Funding**

This work was supported by a research training fellowship funded by the UK National Institute of Health Research to U.R. The report is independent research and the views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

**References**


Table DS1. Overlap of individual PRO definitions with definitions of other PROs.

<table>
<thead>
<tr>
<th>PRO / Reference</th>
<th>Definition</th>
<th>Concepts</th>
<th>Treatment satisfaction</th>
<th>Subjective quality of life</th>
<th>Needs for care</th>
<th>Therapeutic relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment satisfaction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ware et al.15; p. 247</td>
<td>“…a personal evaluation of health care services and providers.”</td>
<td>Multidimensional satisfaction</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>Ruggeri et al.16; p. 265</td>
<td>“…the effect of various factors such as subjects’ previous experience with services, their attitudes toward life, self-esteem and illness behaviour.”</td>
<td>Multidimensional satisfaction</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td><strong>Subjective quality of life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zautra &amp; Goodhard27; p. 4</td>
<td>“Essentially quality of life pertains to the ‘goodness of life.’ We will take the view that this ‘goodness’ resides in the quality of the life experience, both as subjectively evaluated and as objectively determined by an assessment of external conditions.”</td>
<td>Life satisfaction</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Bigelow et al.85; p. 349</td>
<td>“…the social and environmental conditions required to fill basic human needs.”</td>
<td>Role functioning</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Clare et al.28; p. 325</td>
<td>“…the degree of personal satisfaction which he [i.e. the patient] derives from his role performance.”</td>
<td>Role functioning</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Lehman31; p. 52</td>
<td>“…the experience of general well-being as a product of personal characteristics, objective life conditions in various life domains, and satisfaction with life conditions in these various domains.”</td>
<td>Life satisfaction</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>WHOQOL Group86; p. 1403</td>
<td>“…an individual’s 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, expectations, standards and concerns.”</td>
<td>Hybrid satisfaction / role functioning</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Hachey &amp; Mercier87; p. 2</td>
<td>“…is represented by an understanding of patients’ needs and the impact of services on their life.”</td>
<td>Competency</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Degree of overlap: very strong (+++), strong (++), moderate (+), none (−)
<table>
<thead>
<tr>
<th>PRO / Reference</th>
<th>Definition</th>
<th>Concepts</th>
<th>Overlap</th>
<th>Treatment satisfaction</th>
<th>Subjective quality of life</th>
<th>Needs for care</th>
<th>Therapeutic relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subjective quality of life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sintonen(^8^8); p. 328</td>
<td>“…a multidimensional concept that encompasses the physical, emotional and social components associated with illness or treatment.”</td>
<td>Health-related quality of life</td>
<td>+++)</td>
<td>++</td>
<td>++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prieto et al.(^3^5); p. 24</td>
<td>“…encompasses the entire range of human experience, including cultural, physical, psychological, interpersonal, spiritual, financial, political and temporal dimensions.”</td>
<td>Health-related quality of life</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Priebe(^8^9); p. 17</td>
<td>“…represents the person’s appraisal of their objective life conditions, mostly captured by rating scales of satisfaction with life domains and life as a whole.”</td>
<td>Life satisfaction</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Needs for care</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Health Social Services Inspectorate(^2^9); p. 10</td>
<td>“…the requirement of individuals to achieve, maintain or restore an acceptable level of social independence or quality of life, as defined by the particular care agency or authority.”</td>
<td>Supply</td>
<td>–</td>
<td>+++)</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stevens &amp; Gabbay(^\text{24}); p. 201</td>
<td>“The ability to benefit in some way from health care.”</td>
<td>Supply</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phelan et al.(^1^7); p. 589</td>
<td>“For people with a severe mental illness it is appropriate to expand this definition [i.e. from Stevens &amp; Gabbay(^\text{24}); p. 20, see above] to include social care as well as health care.”</td>
<td>Supply</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phelan et al.(^1^7); p. 590</td>
<td>“…a subjective concept, accepting that frequently there will be differing, but equally valid perceptions about the presence or absence of a specific need (Slade, 1994).”</td>
<td>Perceived need</td>
<td>–</td>
<td>++</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Degree of overlap: very strong (+++), strong (++), moderate (+), none (−)
### Table DS1 (cont.)

<table>
<thead>
<tr>
<th>PRO / Reference</th>
<th>Definition</th>
<th>Concepts</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Treatment satisfaction</td>
<td>Subjective quality of life</td>
</tr>
<tr>
<td><strong>Therapeutic relationship</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freud21: p. 139</td>
<td>“…the first aim of treatment is to attach the person of the patient to the person of the therapist.”</td>
<td>Therapeutic</td>
<td>+</td>
</tr>
<tr>
<td>Sterba22: p. 119</td>
<td>“…the need for an alliance between the patient’s mature ego functioning and the working style of the therapist.”</td>
<td>Working alliance</td>
<td>+</td>
</tr>
<tr>
<td>Bordin18: p. 253</td>
<td>“…makes it possible for the patient to accept and follow treatment faithfully.”</td>
<td>Pantheoretical</td>
<td>+</td>
</tr>
<tr>
<td>Pinsof &amp; Catherall20: p. 139</td>
<td>“…that aspect of the relationship between the therapist system and the patient system that pertains to their capacity to mutually invest in, and collaborate on, the therapy.”</td>
<td>Systemic</td>
<td>+</td>
</tr>
<tr>
<td>Priebe &amp; Gruyters19: p. 552</td>
<td>“…the degree to which the patient experiences therapist and therapy as positive and helpful.”</td>
<td>Rogerian</td>
<td>++</td>
</tr>
</tbody>
</table>

*Note: Degree of overlap: very strong (+++), strong (++), moderate (+), none (−)*
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Concept to be measured</th>
<th>Type</th>
<th>Content of domains*</th>
<th>Response options</th>
<th>Number of items</th>
<th>Completion time, min</th>
<th>Psychometric properties assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Satisfaction Questionnaire (PSQ)³⁰</td>
<td>Not reported</td>
<td>Generic</td>
<td>Not reported</td>
<td>Likert scale</td>
<td>9</td>
<td>6</td>
<td>Discriminant validity⁴⁴, concurrent validity⁹¹</td>
</tr>
<tr>
<td>Client Satisfaction Questionnaire (CSQ)⁹²-⁹⁶</td>
<td>Not reported</td>
<td>Generic</td>
<td>Physical surroundings, support staff, kind/type of service, treatment staff, quality of service, quantity of service, outcome of service, procedures</td>
<td>Likert scale</td>
<td>3, 8, 18, 31</td>
<td>5-25 (depending on the version)</td>
<td>Internal consistency⁹²,⁹³,⁹⁵, structural validity⁹²,⁹⁷, discriminant validity⁴⁴, concurrent validity⁹²,⁹³</td>
</tr>
<tr>
<td>Verona Service Satisfaction Scale (VSSS)¹⁶,⁹⁸-¹⁰⁰</td>
<td>Multidimensional satisfaction¹⁵,¹⁶</td>
<td>Generic and treatment-specific parts</td>
<td>Professionals’ skills and behaviours, information, access, efficacy, types of intervention, relatives’ involvement</td>
<td>Likert scale</td>
<td>32, 54, 82</td>
<td>25-40 (depending on the version)</td>
<td>Internal consistency¹⁶,¹⁰⁰, test-retest reliability¹⁰⁰, content validity¹⁵,¹⁰⁰,¹⁰¹, structural validity⁹⁹</td>
</tr>
<tr>
<td>Lund Patient Satisfaction Questionnaire (LPSQ)¹⁰²</td>
<td>Not reported</td>
<td>Generic and treatment-specific parts</td>
<td>Qualities of the staff-patient relationship, ward atmosphere, information about illness and treatment, satisfaction with different types of treatments, qualities of treatment programme, restrictions and compulsory care, accessibility and availability</td>
<td>Likert scale</td>
<td>35, 48</td>
<td>30-40 (depending on the version)</td>
<td>Internal consistency¹⁰², content validity¹⁰²</td>
</tr>
<tr>
<td>Instrument</td>
<td>Concept to be measured</td>
<td>Type</td>
<td>Content of domainsa</td>
<td>Response options</td>
<td>Number of items</td>
<td>Completion time, min</td>
<td>Psychometric properties assessed</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------</td>
<td>---------------------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Client Assessment of Treatment scale (CAT)103</td>
<td>Not reported</td>
<td>Generic and treatment-specific parts</td>
<td>Not reported</td>
<td>Visual analogue scale</td>
<td>7</td>
<td>5</td>
<td>Cross-cultural validity104, predictive validity105, test-retest reliability104, content validity104, convergent validity104</td>
</tr>
<tr>
<td>Psychiatric Care Satisfaction Questionnaire (PCSQ)106,107</td>
<td>Multidimensional satisfaction15,16</td>
<td>Generic</td>
<td>Views of general quality of service, attitudes towards psychiatric doctors in terms of professional qualities and competence, approach and personal qualities, communication and explanation, availability and convenience of contact</td>
<td>Likert scale</td>
<td>18</td>
<td>Not reported</td>
<td>Internal consistency107, test-retest reliability107, content validity107, convergent validity107</td>
</tr>
<tr>
<td>Satisfaction with Antipsychotic Medication (SWAM)108</td>
<td>Not reported</td>
<td>Treatment-specific</td>
<td>Treatment acceptability, medication insight</td>
<td>Likert scale</td>
<td>2</td>
<td>Not reported</td>
<td>Internal consistency108, structural validity on domain scores108</td>
</tr>
<tr>
<td>Medication Satisfaction Questionnaire (MSQ)109</td>
<td>Not reported</td>
<td>Treatment-specific</td>
<td>None</td>
<td>Likert scale</td>
<td>1</td>
<td>1</td>
<td>Test-retest reliability109, convergent validity109, concurrent validity109</td>
</tr>
</tbody>
</table>

a Overlap in content of domains with measures intended to assess different PRO concepts:

- Treatment satisfaction and SQOL: VSSS and CSQ domains overlap with QLQ, QOLI, LOOLP, and SQLS domains
- Treatment satisfaction and needs: VSSS and LPSQ domains overlap with CAN and PNCQ domains
- Treatment satisfaction and therapeutic relationship: CSQ, VSSS and LPSQ domains overlap with STAR, HAS, WAI, and TAS domains
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Concept to be measured</th>
<th>Type</th>
<th>Content of domains</th>
<th>Response options</th>
<th>Number of items</th>
<th>Completion time, min</th>
<th>Psychometric properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Maladjustment Schedule (SMS)(^{28,110})</td>
<td>Role functioning(^{28})</td>
<td>Generic</td>
<td>Housing, social role, economic situation, leisure/social activities, family and domestic relationships, marital situation</td>
<td>Likert scale</td>
<td>48</td>
<td>45</td>
<td>Structural validity(^{28,110})</td>
</tr>
<tr>
<td>Satisfaction with Life Domains Scale (SLDS)(^{111,112})</td>
<td>Life satisfaction(^{27,87}); Competency(^{87})</td>
<td>Generic</td>
<td>Housing, neighbourhood, food, clothing, health, living situation, friends, family, relations with others, day programming, spare time, fun, services and facilities in area, economic situation</td>
<td>Pictorial Likert scale</td>
<td>15</td>
<td>10</td>
<td>Test-retest reliability(^{113,114}), internal consistency(^{114}), concurrent validity(^{112,114})</td>
</tr>
<tr>
<td>Satisfaction With Life (SWL) scale(^{115})</td>
<td>Life satisfaction(^{31})</td>
<td>Generic</td>
<td>Living situation, social relationships, self and present life, work</td>
<td>Likert scale</td>
<td>8, 18, 21</td>
<td>Not reported</td>
<td>Internal consistency(^{115}), structural validity of domain scores(^{115,116})</td>
</tr>
<tr>
<td>Quality of Life Interview (QOLI)(^{31,32})</td>
<td>Life satisfaction(^{31})</td>
<td>Generic</td>
<td>Living situation, family relations, social relations, leisure activities, finances, safety, work and school, health</td>
<td>Likert and visual analogue scales</td>
<td>143</td>
<td>45</td>
<td>Internal consistency(^{31,118}), test-retest reliability(^{31,45}), IRT item information(^{16}), structural validity(^{30,36}), concurrent validity(^{121-123})</td>
</tr>
<tr>
<td>Client Quality of Life Interview (CQLI)(^{124})</td>
<td>Adaptive functioning(^{85,124})</td>
<td>Generic</td>
<td>Leisure and recreation, education and employment, finances, living situation, family, friends and community, ability to cope</td>
<td>Rating scale</td>
<td>65</td>
<td>30</td>
<td>Not reported</td>
</tr>
<tr>
<td>California Well-Being Project Client Interview (CWBPCI)(^{125})</td>
<td>Not reported</td>
<td>Generic</td>
<td>Happiness, health, income, work/achievement, comfort, social life, spiritual life, resources, food, accommodation, sexual life, creativity, basic human freedoms, warmth and intimacy, safety</td>
<td>Likert scale and open-ended questions</td>
<td>151</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Instrument</td>
<td>Concept to be measured</td>
<td>Type</td>
<td>Content of domains&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Response options</td>
<td>Number of items</td>
<td>Completion time, min</td>
<td>Psychometric properties assessed</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-----------------------------------------</td>
<td>----------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>Lancashire Quality of Life Profile (LQOLP)</td>
<td>Life satisfaction&lt;sup&gt;31&lt;/sup&gt;; Adaptive functioning&lt;sup&gt;85,112&lt;/sup&gt;</td>
<td>Generic</td>
<td>Work, finances, social relations, leisure, living situation, safety, family relations, religion, health</td>
<td>Likert scale</td>
<td>100</td>
<td>55</td>
<td>Internal consistency&lt;sup&gt;127-131&lt;/sup&gt;, test-retest reliability&lt;sup&gt;130,131&lt;/sup&gt;, structural validity&lt;sup&gt;131,132&lt;/sup&gt;</td>
</tr>
<tr>
<td>EuroQOL-5D (EQ-5D)&lt;sup&gt;33,34&lt;/sup&gt;</td>
<td>Health-related quality of life&lt;sup&gt;33,35&lt;/sup&gt;</td>
<td>Generic, utility</td>
<td>Mobility, self-care, usual activities, pain/discomfort, anxiety/depression</td>
<td>Likert scale</td>
<td>15</td>
<td>5</td>
<td>Structural validity&lt;sup&gt;35,33&lt;/sup&gt;, concurrent validity&lt;sup&gt;134,135&lt;/sup&gt;, cross-cultural validity&lt;sup&gt;35&lt;/sup&gt;, responsiveness&lt;sup&gt;136-138&lt;/sup&gt;</td>
</tr>
<tr>
<td>Oregon Quality of Life Questionnaire (OQLQ)&lt;sup&gt;85,139&lt;/sup&gt;</td>
<td>Role functioning&lt;sup&gt;85&lt;/sup&gt;</td>
<td>Generic</td>
<td>Psychological distress, psychological well-being, tolerance of stress, total basic need satisfaction, independence, interpersonal interactions, spouse role, social support, work at home, employability, work on the job, meaningful use of time, negative consequences of alcohol use, negative consequences of drug use</td>
<td>Rating and Likert scale</td>
<td>263</td>
<td>45</td>
<td>Internal consistency&lt;sup&gt;139&lt;/sup&gt;, concurrent validity&lt;sup&gt;139&lt;/sup&gt;</td>
</tr>
<tr>
<td>Medical Outcomes Study 36 Item Short Form Health Survey (SF-36)&lt;sup&gt;140-142&lt;/sup&gt;</td>
<td>Health-related quality of life&lt;sup&gt;140&lt;/sup&gt;</td>
<td>Generic</td>
<td>Physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, mental health</td>
<td>Likert scale</td>
<td>12, 36</td>
<td>5-10</td>
<td>Internal consistency&lt;sup&gt;143,145&lt;/sup&gt;, test-retest reliability&lt;sup&gt;147&lt;/sup&gt;, structural validity&lt;sup&gt;143,145&lt;/sup&gt;, concurrent validity&lt;sup&gt;144&lt;/sup&gt;</td>
</tr>
<tr>
<td>15D&lt;sup&gt;88,148&lt;/sup&gt;</td>
<td>Health-related quality of life&lt;sup&gt;88,148&lt;/sup&gt;</td>
<td>Generic, utility</td>
<td>Mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort and symptoms, depression, distress, vitality, sexual activity</td>
<td>Rating and ratio scales</td>
<td>15</td>
<td>5-10</td>
<td>Responsiveness&lt;sup&gt;137&lt;/sup&gt;</td>
</tr>
<tr>
<td>Subjective Well-Being under Neuroleptics Scale (SWN)&lt;sup&gt;140-152&lt;/sup&gt;</td>
<td>Subjective effects of neuroleptics</td>
<td>Disease-specific</td>
<td>Mental functioning, social integration, emotional regulation, physical functioning, self-control</td>
<td>Likert scale</td>
<td>5, 20, 38</td>
<td>(depending on the version)</td>
<td>Internal consistency&lt;sup&gt;149,150,153,154&lt;/sup&gt;, structural validity&lt;sup&gt;155&lt;/sup&gt;, concurrent validity&lt;sup&gt;149,150&lt;/sup&gt;</td>
</tr>
<tr>
<td>Instrument</td>
<td>Concept to be measured</td>
<td>Type</td>
<td>Content of domains&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Response options</td>
<td>Number of items</td>
<td>Completion time, min</td>
<td>Psychometric properties assessed</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------</td>
<td>------</td>
<td>-------------------------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Health Utilities Index Mark 3 (HUI3)&lt;sup&gt;156,157&lt;/sup&gt;</td>
<td>Health-related quality of life&lt;sup&gt;157&lt;/sup&gt;</td>
<td>Generic, utility</td>
<td>Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain</td>
<td>Likert scale</td>
<td>8</td>
<td>5</td>
<td>Convergent validity&lt;sup&gt;158&lt;/sup&gt;</td>
</tr>
<tr>
<td>Quality of Life Questionnaire (QLQ)&lt;sup&gt;159&lt;/sup&gt;</td>
<td>Life satisfaction&lt;sup&gt;31&lt;/sup&gt;</td>
<td>Generic</td>
<td>Living situation, finances, family, social life, leisure, living situation, health, access to medical care</td>
<td>Likert scale</td>
<td>24</td>
<td>10</td>
<td>Internal consistency&lt;sup&gt;159&lt;/sup&gt;</td>
</tr>
<tr>
<td>World Health Organization Quality of Life Instrument (WHOQOL)&lt;sup&gt;86,160,161&lt;/sup&gt;</td>
<td>Hybrid satisfaction/role functioning&lt;sup&gt;86&lt;/sup&gt;</td>
<td>Generic</td>
<td>Physical, psychological, social relationships, environment</td>
<td>Likert-type scale</td>
<td>26, 100</td>
<td>Not reported</td>
<td>Internal consistency&lt;sup&gt;161&lt;/sup&gt;, structural validity&lt;sup&gt;159&lt;/sup&gt;, convergent validity&lt;sup&gt;138&lt;/sup&gt;, responsiveness&lt;sup&gt;138&lt;/sup&gt;, concurrent validity&lt;sup&gt;138&lt;/sup&gt;</td>
</tr>
<tr>
<td>Manchester Short Assessment of Quality of Life (MANSA)&lt;sup&gt;166&lt;/sup&gt;</td>
<td>Life satisfaction&lt;sup&gt;31&lt;/sup&gt;</td>
<td>Generic</td>
<td>Work, finances, friends, leisure, accommodation, people that the patient lives with, safety, sex life, family relations, physical health, mental health</td>
<td>Likert scale</td>
<td>25</td>
<td>Not reported</td>
<td>Convergent validity&lt;sup&gt;166&lt;/sup&gt;, discriminant validity&lt;sup&gt;44&lt;/sup&gt;</td>
</tr>
<tr>
<td>Schizophrenia Quality of Life Scale (SQLS)&lt;sup&gt;47&lt;/sup&gt;</td>
<td>Not reported</td>
<td>Disease-specific</td>
<td>Psychosocial, motivation and energy, symptoms and side effects</td>
<td>Likert-type scale</td>
<td>30</td>
<td>5-10</td>
<td>Internal consistency&lt;sup&gt;47,167&lt;/sup&gt;, structural validity&lt;sup&gt;168,167,169&lt;/sup&gt;, concurrent validity&lt;sup&gt;170&lt;/sup&gt;</td>
</tr>
<tr>
<td>Impact of Weight on Quality of Life-Lite (IWQOL-Lite)&lt;sup&gt;172&lt;/sup&gt;</td>
<td>Weight-related quality of life&lt;sup&gt;172&lt;/sup&gt;</td>
<td>Disease-specific</td>
<td>Physical function, self-esteem, sexual life, work, and public distress</td>
<td>Likert scale</td>
<td>31</td>
<td>Not reported</td>
<td>Internal consistency&lt;sup&gt;172,173&lt;/sup&gt;, test-retest reliability&lt;sup&gt;172,173&lt;/sup&gt;, concurrent validity&lt;sup&gt;172,173&lt;/sup&gt;</td>
</tr>
<tr>
<td>Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q)&lt;sup&gt;48,174&lt;/sup&gt;</td>
<td>Enjoyment / satisfaction&lt;sup&gt;48&lt;/sup&gt;</td>
<td>Generic</td>
<td>Physical health, subjective feelings, leisure activities, social relationships, general activities, satisfaction with medication, life satisfaction</td>
<td>Likert scale</td>
<td>93, 18</td>
<td>10-45, (depending on the version)</td>
<td>Internal consistency&lt;sup&gt;49,174&lt;/sup&gt;, test-retest reliability&lt;sup&gt;174&lt;/sup&gt;, convergent validity&lt;sup&gt;24,178&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Overlap in content of domains with measures intended to assess different PRO concepts:

SQOL and treatment satisfaction: QLQ, QOLI, LQOLP, and SQLS domains overlap with VSSS and CSQ domains

SQOL and needs: SMS, SLDS, QOLI, CQLI, CWBPCI, LQOLP, OQLQ, QLQ, WHOQOL, and MANSA domains overlap with CAN and PNCQ domains
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Concept to be measured</th>
<th>Type</th>
<th>Content of domains</th>
<th>Response options</th>
<th>Number of items</th>
<th>Completion time, min</th>
<th>Psychometric properties assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camberwell Assessment of Need (CAN)(^{17})</td>
<td>Supply(^{17} /) Perceived need(^{17})</td>
<td>Generic / disease-specific</td>
<td>Accommodation, food, looking after the home, self care, daytime activities, physical health, psychotic symptoms, information, psychological distress, safety to self, safety to others, alcohol, drugs, company, intimate relationships, sexual expression, childcare, basic education, telephone, transport, money and benefits</td>
<td>Checklist</td>
<td>22, 66</td>
<td>10-30 (depending on the version)</td>
<td>Internal consistency(^{17,179-181}), test-retest reliability(^{17,179-182}), face validity(^{17,179,181}), content validity(^{17,179,181}), concurrent validity(^{17,179,181}), discriminant validity(^{44}), predictive validity(^{183})</td>
</tr>
<tr>
<td>Perceived Need for Care Questionnaire (PNCQ)(^{184})</td>
<td>Not reported</td>
<td>Generic</td>
<td>Drugs, information, psychotherapy, social intervention, skills training</td>
<td>Likert-type scale</td>
<td>8</td>
<td>2</td>
<td>Convergent validity(^{184})</td>
</tr>
<tr>
<td>Two-Way Communication Checklist (2-COM)(^{185,186})</td>
<td>Not reported</td>
<td>Generic / disease-specific</td>
<td>Not reported</td>
<td>Likert-type scale</td>
<td>19</td>
<td>10-15</td>
<td>Internal consistency(^{185,186}), test-retest reliability(^{185,186}), face validity(^{185,186}), concurrent validity(^{185,186,187})</td>
</tr>
<tr>
<td>Needs Assessment Scale (NAS)(^{37})</td>
<td>Not reported</td>
<td>Generic / disease-specific</td>
<td>Accommodation, food, looking after the home, self care, daytime activities, physical health, psychotic symptoms, information, psychological distress, safety to self, safety to others, alcohol, drugs, company, intimate relationships, sexual expression, childcare, basic education, telephone, transport, money and benefits</td>
<td>Likert-type scale</td>
<td>44</td>
<td>Not reported</td>
<td>Structural validity of subscales(^{37}), concurrent validity(^{37})</td>
</tr>
</tbody>
</table>

\(^{a}\) Overlap in content of domains with measures intended to assess different PRO concepts:

*Needs and treatment satisfaction*: CAN and PNCQ domains overlap with VSSS and LPSQ domains

*Needs and SQOL*: CAN and PNCQ domains overlap with SMS, SLDS, QOLL, CQLI, CWBPCI, LQOLP, OQLQ, QLQ, WHOQOL, and MANSA domains
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Concept to be measured</th>
<th>Type</th>
<th>Content of domains</th>
<th>Response options</th>
<th>Number of items</th>
<th>Completion time, min</th>
<th>Psychometric properties assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrative Psychotherapy Alliance Scale (IPAS)(^{20})</td>
<td>Pantheoretical(^{18})</td>
<td>Generic</td>
<td>Goals, bonds, tasks</td>
<td>Likert scale</td>
<td>25</td>
<td>10</td>
<td>Not reported</td>
</tr>
<tr>
<td>Working Alliance Inventory (WAI)(^{18})</td>
<td>Pantheoretical(^{18})</td>
<td>Generic</td>
<td>Goals, bonds, tasks</td>
<td>Likert scale</td>
<td>36, 12</td>
<td>10</td>
<td>Internal consistency(^{189}), concurrent validity(^{189}), structural validity(^{190}), predictive validity(^{189})</td>
</tr>
<tr>
<td>Therapeutic Alliance Scale (TAS)(^{19})</td>
<td>Not reported</td>
<td>Generic</td>
<td>Confidence, quality of the relationship</td>
<td>Likert scale</td>
<td>2</td>
<td>1</td>
<td>Not reported</td>
</tr>
<tr>
<td>Helping Alliance Scale (HAS)(^{19})</td>
<td>Rogerian(^{19})</td>
<td>Generic</td>
<td>Feeling understood, feeling respected, receiving the right care</td>
<td>Visual Analogue Scale</td>
<td>6</td>
<td>5</td>
<td>Convergent validity(^{193}), concurrent validity(^{193}), discriminant validity(^{44})</td>
</tr>
<tr>
<td>California Psychotherapy Alliance Scales (CALPAS)(^{194})</td>
<td>Hybrid(^{18,21,22})</td>
<td>Generic</td>
<td>Therapeutic alliance, working alliance, therapists’ understanding and involvement, patient-therapist agreement on goals and strategies</td>
<td>Likert scale</td>
<td>31, 24</td>
<td>10</td>
<td>Not reported</td>
</tr>
<tr>
<td>Scale to Assess the Therapeutic Relationship (STAR)(^{195})</td>
<td>Not reported</td>
<td>Generic</td>
<td>Positive collaboration, positive clinician input, non-supportive clinician input</td>
<td>Likert scale</td>
<td>12</td>
<td>5-10</td>
<td>Test-retest reliability(^{195}), structural validity of subscales(^{195})</td>
</tr>
</tbody>
</table>

\(^a\) Overlap in content of domains with measures intended to assess different PRO concepts:

*Therapeutic relationship and treatment satisfaction:* STAR, HAS, WAI, and TAS domains overlap with CSQ, VSSS and LPSQ domains


Ruggeri M, Dall’Agnola R. The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients, relatives and professionals. Psychol Med 1993; 23: 511-23.


Measuring patient-reported outcomes in psychosis: conceptual and methodological review
Ulrich Reininghaus and Stefan Priebe
Access the most recent version at DOI: 10.1192/bjp.bp.111.107615