Symptom rating scales and outcome in schizophrenia

ANN M. MORTIMER

Background Symptom rating scales are now well established in schizophrenia research but their scores are not the same as outcome.

Aims To appraise the usefulness of symptom rating scales in evaluating the outcome of people with schizophrenia.

Method Literature on the use of the Brief Psychiatric Rating Scale (BPRS) the Positive and Negative Syndrome Scale (PANSS) and the Clinical Global Impression (CGI) in schizophrenia research was studied.

Results Scales were designed to make diagnoses, to categorise patients, syndromes or both, and to demonstrate antipsychotic efficacy, as well as to measure outcome. There is much redundancy both between and within scales. Early work suggests limited concurrent validity with external outcome variables. Data are at best ordinal and there are particular difficulties in equating outcome with percentage changes in scores. The concept of remission, which uses absolute item score thresholds with a duration criterion, is a promising outcome measure.

Conclusions Symptom rating scale scores can only comprise a limited part of outcome measurement. Standardised remission criteria may present advantages in outcome research.

Outcome measures are important in schizophrenia because we need to identify whether outcomes are modified by the medications and psychosocial interventions which we offer. Leaving aside social cultural and environmental factors, before the antipsychotic era it is unlikely that outcome was influenced by anything other than the intrinsic nature and severity of the schizophrenic illness. Providing basic nursing care and protection probably influenced negative outcomes to some extent.

Outcome is not a unitary construct defined simply by lack of symptoms: personal and social function, cognition and quality of life must be of substantial relevance. Other aspects such as economic outcome, although important to commissioners and providers of services, might be of limited consequence to clinicians and patients, who naturally focus on professional and consumer (satisfaction) standpoints respectively. Hence, outcome evaluation applied to services differs from that applied to patients.

Symptom rating scales in schizophrenia were not initially designed to assess the efficacy of antipsychotic drug treatments. Nevertheless, they have been used in this role more than any other. This is not surprising as antipsychotic drugs are used primarily to control patients' symptoms; the underlying neuroscience is consistent with this, and not with any direct therapeutic effects on cognition, personal and social function, or quality of life (unless mediated by symptom control). Although such distal effects have been proposed, there are numerous independent variables which influence these aspects of outcome (e.g. upbringing, premorbid personality and adjustment, intellect and mood, social circumstances and availability of a support network). Furthermore, it has been proposed that antipsychotic drugs, particularly conventional antipsychotics, have little effect on negative symptoms of schizophrenia. Negative symptoms are one of the most clinically important targets, and overlap with cognition and function (Mortimer & Spence, 2001).

SYMPTOM RATINGS AS OUTCOME MEASURES

Although there is evidence that changes in distinct psychopathological dimensions differentially influence broader aspects of outcome (Van Os *et al*, 1996) it is now accepted that fixed factors such as duration of untreated psychosis, gender, age of onset and family psychiatric history make a substantial contribution (Murray & Van Os, 1998). Symptom rating scales can be viewed as quantifying the skilled clinician's judgement of current psychopathology, and change over time. The worth of routine use of such rating scales in ordinary clinical practice is the subject of continuing debate; the clinician makes an initial, comprehensive assessment of the patient, and reviews this as treatment proceeds and the final outcome becomes clearer. The added value of a highly structured approach can be questioned in a clinical review of an individual patient's progress. Most patients manifest only a minority of the range of possible symptoms and generally do not develop too many new symptoms during treatment. In routine practice, symptom scales are perhaps little more than a formalised guide to what the clinician should be doing already. They have specific utility in training junior staff in the full range of psychopathology they are likely to encounter, and the finer points of mental state examination. Repeated scores, represented graphically, may have some utility in communicating a patient's progress to other clinicians. In research, symptom rating scales in schizophrenia will inform the investigator what is the nature and 'volume' of symptoms experienced by the patient, and the magnitude of any change over time.

Limitations

Symptom rating scale data can never be anything more than ordinal; the overall total of symptom item scores will often lump together categorical data, containing symptoms associated in clusters, such as the positive, negative and disorganisation syndromes. Specific syndrome scores derived from scales may have more utility than the total score regarding an overall perspective. Current thinking includes that schizophrenia syndromes may comprise positive

(disorganisation and reality distortion) and negative categories, with non-negative affective symptoms (mostly depressive) in a significant minority of patients. Consequently three or four syndrome scores in the context of a defined range may give a reasonable 'snapshot' of a patient's current clinical status. Such quantification may inform judgement regarding aetiology, treatment and prognosis (Van Os *et al*, 1996). For example, negative symptoms are known to have adverse consequences for personal and social function and cognition (Rocca *et al*, 2005). By contrast, even extensive, but isolated, reality distortion may generate minimal functional consequence, whereas disorganisation syndrome is usually very disruptive (Schuldberg *et al*, 1999). Depression may arise from several sources, with varying outcome (Emsley *et al*, 1999). Such data have implications for treatment interventions. The Clinical Global Impression–Schizophrenia scale (CGI–SCH; Haro *et al*, 2003) represents, conceivably, a step in this direction although its positive, negative, depression and cognitive scores are rated according to judgement of severity rather than from items comprising these syndromes.

The value of symptom item or even syndrome score totals *per se* is increasingly questioned in the determination of outcome status. A more patient-centred definition of outcome, stressing personal and social function, is often viewed as more practical than the presence or absence of esoteric phenomena (symptoms), which may have little bearing on subjective experience or uptake of healthcare. Influential work has attempted to explore the meaning and consequences of delusions and hallucinations for patients (Chadwick & Birchwood, 1995), but scales derived from this work are not in widespread use outside the research setting. Self-administered symptom scales have been developed (Hamera *et al*, 1996) but again these have not found wide usage, in contrast to the emphasis on patient-rated quality of life as an outcome. Clinicians increasingly seek treatment outcomes such as degree of independent living, time to discontinuation of medication, and time to relapse and rehospitalisation rather than changes in symptom rating scale scores (Tiihonen *et al*, 2006).

Concurrent validity

The question remains whether any rating scale (or factorial components of it) demonstrates sufficient concurrent validity to predict these external outcome variables. Operational definitions of remission may achieve this. These consist of multiple item threshold rather than factorial scores, with the addition of a duration condition. In the absence of concurrent validity with other outcome measures, symptom rating scales can only constitute a small part of the appraisal of overall outcome. Symptom rating scales will answer the question 'Did the antipsychotic drug work on this patient's symptoms?' as opposed to 'What is this patient's outcome?' Marshall *et al*, 2000 emphasise that the use of unpublished rating scales in controlled trials is associated with consistent claims of superiority of new treatments and that familiar, well-validated scales may give a more accurate answer.

THE BIG THREE

Three symptom rating scales have dominated the field of schizophrenia research and, in particular, studies of antipsychotic efficacy. With the admonition of Marshall (Marshall *et al*, 2000) in mind they will be dealt with in some detail here.

Brief Psychiatric Rating Scale

The Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962) is a one-page, 16- or 18-item rating scale which was developed more than 40 years ago. It assesses a range of psychotic and affective symptoms rated from both observation of the patient and the patient's own report. The original purpose of the BPRS was the rapid evaluation of clinical change irrespective of origin (e.g. natural remission or treatment response) in the broad range of psychiatric patients, not just those with schizophrenia. It was not, therefore, specifically designed as an outcome measure; the authors hoped that the scale would develop into a diagnostic instrument, which they considered of greater long-term value than detecting change. Standard definitions of outcome were developed later, e.g. 'consumer outcome is the effect on a patient's health status attributable to an intervention by a health professional or health service' (Andrews *et al*, 1994). Even so, the authors later stated that the BPRS was designed to fill a special need in clinical psychopharmacology research, at the inception of the Early Clinical Drug Evaluation Units of the National Institute of Mental Health in the USA (Overall & Gorham, 1988).

Extent of use and adaptation

The BPRS has perhaps been used more extensively than any other symptom rating scale, in many diagnostic groups and for a wide range of purposes. It is highly sensitive to change, and excellent interrater reliability can be achieved with training and a standard interview procedure (Overall & Rhoades, 1982). As well as the evaluation of efficacy of several classes of psychotropic medication (Hedlund & Vieweg, 1980; Overall & Rhoades 1982; Perry *et al*, 1997; Hamilton *et al*, 1998), the BPRS has been used extensively to compare diagnostic concepts internationally and in epidemiological studies (Delmonte *et al*, 1970; Engelsmann & Formankova, 1967; Engelsmann *et al*, 1970; Overall & Beller, 1984). It has been translated into many languages and frequently modified for specific purposes, including for use with children (Overall & Pfefferbaum, 1982; Emslie *et al*, 1997). It has been expanded to 24 items to make it more comprehensive in the area of psychotic and affective symptoms, with items on bizarre behaviour, suicidality, self-neglect, elevated mood, distractability and motor hyperactivity (Ventura *et al*, 2000). The BPRS has been demonstrated as reliable for use by nursing staff, increasing its utility (McGorry *et al*, 1988). Most adaptations of the BPRS use one of two scoring versions for each item (either a 0- to 3-point or a 0- to 7-point scale).

Limitations

The factor structure of BPRS responses depends upon the characteristics of the patient group under study, and the version being used. The BPRS was, until the advent of the Positive and Negative Syndrome Scale (PANSS; Kay *et al*, 1987) which itself is partially derived from the BPRS, the most widely used scale in schizophrenia research. This reflected its broad coverage of typical schizophrenia phenomena in the positive, negative and disorganisation categories. However, its coverage of the negative syndrome has been criticised; there are only three negative syndrome items, and it has been suggested that a more extensive scale is necessary for sensitivity to change (Eckert *et al*, 1996).

The authors themselves were dismissive of the use of their scale to determine differences between specific symptoms or

syndromes during treatment, stating that 'Although psychiatric symptomatology is multidimensional, the difference between pre-treatment pathology and post-treatment pathology (or lack of it) can be represented by a single dimension spanning the multivariate space' (Overall & Gorham, 1988). Despite this, with the assistance of 20 psychiatrists, they gave 13 different weights to each item according to diagnosis, in order to increase or reduce the relevance of treatment effects to the total score. For instance, the score on item 8, 'grandiosity', would be multiplied by a 0 in a patient with depression and by 3 in a patient with paranoia. This complex and somewhat arbitrary scoring system appears never to have been taken up.

Clinical Global Impression

The CGI is not strictly a symptom rating scale but is included because of its wide use, influence and the recent development of forms specific to the schizophrenia syndromes (CGI–SCH). The original version is a simple instrument which rates the overall severity of any mental disorder (Guy, 1976). This is rated entirely according to clinical judgement in routine professional practice, on a scale for the overall current severity of symptoms from 1 (healthy, not ill) to 7 (among the most severely ill). There is also a 7-point scale for global improvement (usually from baseline to the current condition), rating from 1 (very much improved) to 7 (very much worse). The CGI has been used in several efficacy and effectiveness studies in schizophrenia, is sensitive to change and correlates well with changes assessed with more complex scales (Haro et al, 2003; Leucht & Engel, 2006; Leucht et al, 2006; Rabinowitz et al 2006).

The main criticism levelled at the CGI, that it lacks standard definitions (Beneke & Rasmus, 1992), reflects what many consider its main strength – the use of an adequate level of clinical judgement. Its brevity, utility and appeal to clinical commonsense have ensured its continued use over many more complex rating scales. The CGI has been adapted for the assessment of bipolar affective disorder (CGI–BP) and schizophrenia (Spearing et al, 1997; Haro et al, 2003). The CGI–SCH has demonstrated good reliability and validity in the evaluation of severity of positive, negative, depressive and cognitive symptoms, and is recommended for both research and clinical practice.

Positive And Negative Syndrome Scale

The PANSS (Kay et al, 1987, 1988, 1989) originated from a growing need to reduce the heterogeneity of what was known about schizophrenia. Crow's (Crow, 1980) positive–negative dichotomy presented a promising theoretical model for explaining and understanding variability in the aetiology of schizophrenia, treatment and prognosis. However, attempts to utilise the model in practice met with inconsistent results (Andreasen, 1982; Andreasen & Olsen, 1982; Pogue-Geile & Harrow, 1984; Lindenmayer et al, 1986), and it was suggested that this might be because of the lack of a comprehensive rating scale for positive and negative symptoms that was feasible, accurate, well validated, reliable, sensitive and standardised. The PANSS, therefore, was not developed to assess outcome per se, or even the results of treatment interventions.

Nature and scoring

The PANSS is a 30-item 7-point (1–7) rating scale which amalgamated the 18-item BPRS and 12 items from the Psychopathology Rating Schedule (Singh & Kay, 1975). The items were precisely defined, as were anchor points for the numerical rating of each item. The PANSS was divided into positive, negative and general psychopathology sub-scales (a 'manic' sub-scale was later derived; Lindenmayer et al, 2004) and trialled on over 100 well-characterised patients with chronic illness. Sub-scale scores were shown to be normally distributed and independent of each other; they were robust to the effects of mood, chronicity, medication side-effects and cognition. The PANSS was furthermore sensitive and specific regarding pharmacological manipulation of the levels of both positive and negative symptoms in patients with schizophrenia. The validity of its sub-scales was confirmed in an exploration of a classification of patients by predominant symptom class. Sub-scale scores were associated with a number of clinical, treatment and cognitive variables, including premorbid adjustment (Krauss et al, 1998), but not outcome. One of the strengths claimed for the PANSS is consistency in scoring individual patients over time and illness course. A potentially confusing feature of the PANSS, however, is that even those without any mental ill health will score 30. In effect, this means that 30 must be subtracted from the

patient's score in order to gain a meaningful understanding.

Correlations and factors

Several studies have sought correlations between PANSS total and sub-scale scores, and other aspects of the illness, to demonstrate concurrent validity. Other aspects have included ventricular enlargement and cortical atrophy (d'Amato et al, 1992), work performance (Bell et al. 1992), neuropsychological impairment (Bell et al, 1994; Liu et al, 1997; Mass et al, 2000; Bozikas et al, 2004; Good et al, 2004; Ritsner et al, 2006) and violent behaviour (Steinert et al, 2000). Overall these findings appear not to be sufficiently convincing as to be of clinical use, and PANSS scores have generally not been used as proxy variables. For example, when PANSS 'cognitive' items were used to predict global cognitive function 66% of the variance was unexplained, suggesting that the PANSS lacked sensitivity and specificity in this regard (Good et al, 2004). This approach appears not to have generated further research hypotheses.

Factorial validity (the nature and purity of the syndromal components of the scale) is essential to the success of investigations utilising sub-scale scores. There are many reports on the factor (syndrome) structure of PANSS items, with much controversy over whether data best fit a three-, four-, five- or even six-factor solution (Peralta & Cuesta, 1994; Lindenmayer et al, 1994; Wolthaus et al, 2000; Fresan et al, 2005; White, 2005; Van den Oord et al, 2006). The simplest factor solutions comprise a syndrome made up of negative symptom items (psychomotor poverty syndrome), a syndrome made up of delusions and hallucinations (reality distortion syndrome) and a syndrome made up of thought disorder and inappropriate affect symptom items (disorganisation syndrome). Although several five-factor models have been proposed, none has been validated by confirmatory factor analysis (van der Gaag et al, 2006a). This might reflect the ambiguous definitions of some symptom items, such as lack of judgement and insight, which have more than one cause in schizophrenia.

Another complication is that the depression sub-scale (unlike the Calgary Depression Scale; Addington et al, 1992) is unable to distinguish between depression, negative symptoms and extrapyramidal side-effects (Collins et al, 1996). Negative factor scores have been found to correlate

with an independent depression rating instrument (Montgomery Åsberg Depression Rating Scale), although depression factor scores did as well (Wolthaus *et al*, 2000). The loading of single items by multiple causes, which was suggested in another study (Van den Oord *et al*, 2006) was confirmed in a statistically novel analysis (van der Gaag *et al*, 2006*b*).

Only if syndromes possess concurrent validity with other aspects of schizophrenia such as cognitive impairment and poor social function, and furthermore fit explanatory data, can they represent clinical reality. The implication for the rating scale is that items which load on more than one factor must be replaced by two or more items, each of which load on a single factor, which results in lengthier scales. The alternative is losing data through deletion of such items. Poor fit suggests that correlations between syndrome scores and other illness variables under investigation, including outcome, might be unreliable.

MEANING OF SYMPTOM RATING SCALE SCORES

The existence of apparently rival rating scales can be confusing when they purport to measure the same thing. Despite the caveats regarding factorial purity which have been repeatedly addressed in the case of the PANSS, there appears to be much redundancy both within and between rating scales. For example, there are high correlations between positive and negative syndrome scores on the PANSS, and Andreasen's Schedule for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984*a*) and Schedule for the Assessment of Negative Symptoms (SANS; Andreasen, 1984*a*,*b*; Norman *et al*, 1996). The negative symptoms of the PANSS and BPRS, and the SANS all measure, mostly, affective flattening rather than the full range of negative symptom phenomena (Welham *et al*, 1999). The much shorter and quicker CGI scales were just as good as the BPRS in discriminating between the effects of antipsychotic drugs (Leucht & Engel, 2006) despite having been criticised on semantic, logical and statistical grounds (Beneke & Rasmus, 1992). The development of the CGI–SCH scale suggests that investment in less complex rating instruments is gathering pace for rating severity and treatment response in routine clinical practice (Haro *et al*, 2003).

Even in randomised placebo-controlled trials for licensing purposes, the use of changes in rating scale scores may lack good face validity. Many trials evaluate clinical response as a percentage change in scores over the treatment period. Equating a 20% improvement in symptoms with response follows the study of Kane *et al* (1988) which compared clozapine and chlorpromazine in treatment-resistant patients with severe illness. This relatively low percentage reflects the fact that in patients with severe illness even a fairly small attenuation of symptoms might be clinically valuable. The 20% definition of response might not, however, be generalisable to the majority of acute trials with non-resistant patients. Relying on percentage point change to indicate recovery ignores the importance of baseline levels. A 20% reduction of a PANSS score of 100 is double a 20% reduction of a PANSS score of 50, yet both might be recorded as a 'clinical response'. The patient with a baseline PANSS score of 100 would, although fulfilling criteria for response with a score of 80, remain severely ill, (albeit noticeably less so), whereas the patient with a baseline score of 50 would remain mildly ill with a score of 40 and perhaps not even be noticeably different.

Concurrent validity

Leucht *et al*, 2005*a* addressed the issue of what rating scale scores mean in clinical terms. They used an equating procedure to anchor BPRS scores to CGI categories (both severity and improvement) across seven drug trials which used both scales in patients with acute schizophrenia. Clinician-rated 'minimal improvement' on the CGI equated to a 30% improvement on the BPRS (substantially greater than the generally accepted standard for response). 'Much improvement' after 4 weeks of treatment equated to a fall in the BPRS score of almost 58% (Table 1). In addition they found that clinicians used only a small part of the BPRS score range of 18–126: patients with minimum illness on the CGI scored 31, those with moderate illnesss scored 41 and those with severe illness 53. This is probably because patients are only assessed on a minority of the items and upon most they are scored zero.

Using the same approach with the PANSS (Leucht *et al*, 2005*b*) they found that 'mildly ill', 'moderately ill', 'markedly ill' and 'severely ill' according to the CGI equated to total PANSS scores of 58, 75, 95 and 116 respectively (Table 2). At 6 weeks, to achieve CGI ratings of 'minimally improved' and 'much improved' the PANSS decrements were 28% and 53%. The authors suggested that response ought to be defined as a 50% improvement in PANSS score, although in treatment-resistant groups a decrement of 25% might suffice.

A later study (Leucht *et al*, 2006) compared the PANSS and BPRS with each other and with the CGI and replicated the findings overall, emphasising that smaller absolute score reductions equated to perception of improvement in patients with severe illness compared with those with mild illness (Table 3). For a reduction of 1 point on the CGI Severity of Illness scale there were decreases of 15 and 10 on the PANSS and BPRS respectively.

A similar study (Cramer *et al*, 2001) found that clinician-rated 'improved' and 'much better' patients had PANSS scores lowered by 21 and 45% respectively. Quality of life scores were also increased by similar degrees (26 and 50%). This is consistent with the Leucht *et al* (2006) study, and perhaps demonstrates some concurrent validity of the PANSS with subjective quality of life as an outcome. A further report indicated that a decrement of 20% on the PANSS equated to a 1-point severity decrease on the CGI–SCH (Rabinowitz *et al*, 2006).

Table I Clinical implications of BPRS scores

Severity of illness	Corresponding BPRS score	CGI Global Improvement	Corresponding BPRS reduction at I, 2 and 4 weeks, %
Mildly ill	31	Minimally improved	24, 27 and 30
Moderately ill	41	Much improved	40, 53 and 58
Markedly ill	53	Very much improved	71, 79 and 85

BPRS, Brief Psychiatric Rating Scale; CGI, Clinical Global Impression.

Table 2 Clinical implications of PANSS scores

CGI Severity of Illness	Corresponding PANSS score	CGI Global Improvement	Corresponding PANSS reduction at 1, 2, 4 and 6 weeks, %
Mildly ill	58	Minimally improved	19, 23, 26 and 28
Moderately ill	75	Much improved	40, 45, 51 and 53
Markedly ill	95	Very much improved	71, 73, 82 and 81
Severely ill	116	–	–

CGI, Clinical Global Impression; PANSS, Positive and Negative Syndrome Scale.

Table 3 CGI Global Improvement in relation to absolute reductions in PANSS and BPRS scores

Global CGI improvement	Reduction in PANSS score (range)	Reduction in BPRS score (range)
Minimally improved	15 (12–18)	10 (8–11)
Much improved	33 (30–36)	20 (19–22)

BPRS, Brief Psychiatric Rating Scale; PANSS, Positive and Negative Syndrome Scale.

REMISSION

These practical difficulties in the use of symptom rating scales to evaluate outcome in treatment trials have contributed to the recent development of the concept of remission in schizophrenia. Response to treatment focuses on short-term improvements and gives little guidance to clinicians regarding long-term management. In general medicine, remission implies a low level of symptoms but with functional recovery. A number of disparate definitions of remission in schizophrenia have been constructed (Leucht & Lasser 2006). A standard definition, it is argued, is potentially useful: it is realistic and establishes a meaningful treatment goal. Although a useable measure will not include cognition, personal and social function because of difficulties in measurement, there is some evidence that concepts of remission based on symptoms and duration are indeed associated with such consequential aspects of patients' well-being (Birsoz et al, 2006).

The Remission in Schizophrenia Working Group was convened in April 2003 to develop a consensus definition of remission in schizophrenia (Andreasen et al, 2005). Taking precedents in physical medicine and affective disorder, remission should be defined as low or mild symptom levels (which by definition do not influence behaviour) and which should last for a minimum, defined duration. Such a standardised definition, unlike several previous published definitions, could be applied across treatment studies and would permit immediate, transparent comparison. This approach does, however, require attention to levels of baseline severity across studies.

The Working Group aimed to map the chosen remission symptoms, which had to be rated mild or less, onto the three best validated syndromes of schizophrenia (reality distortion, disorganisation and negative symptoms) and the five DSM–IV criteria for schizophrenia (delusions, hallucinations, disorganised speech, disorganised or catatonic behaviour, negative symptoms, American Psychiatric Association, 1994). They picked appropriate items from the BPRS, the PANSS the SAPS and the SANS (Table 4).

The BPRS, with limited coverage of negative symptoms, was perhaps less useful in determining remission. The Working Group set 6 months as the minimum duration of symptoms remaining mild for the patient to qualify for remitted status.

Use of remission criteria

Remission is already being used in attempts to test efficacy of drugs in 'head to head' comparisons by re-analysing existing data (Sethuraman et al, 2005). A study of stable patients using PANSS-based remission criteria demonstrated that nearly 70% were not in remission; 20% achieved remission when switched to depot treatment and 85% of those already in remission remained so a year later on depot (Lasser et al, 2005). Application of the criteria to data from other published studies produced similar findings (Gharabawi et al, 2005; Kissling et al, 2005). In all studies remission was associated with PANSS total and subtotal scores, CGI–SCH scores, functioning and quality of life. Moreover, an analysis of six clinical trials comparing two definitions, one PANSS based and the other BPRS/CGI based, found that achievement of remission using either definition was associated with better quality of life (Dunayevich et al, 2006). This was particularly so if remission was sustained. Nevertheless, total BPRS change score still contributed the greatest part of the variance in quality of life.

Two reviews of the Working Group remission criteria (Nasrallah, 2006; Van Os et al, 2006) proposed that the definition was conceptually viable and feasible in both clinical trials and clinical practice. Both reviews considered that the use of remission criteria would raise clinical expectations and drive clinical services to achieve and document better outcomes. In clinical trials, the concept should improve the quality of methodology and data reporting, while extending its relevance to cognition and functional outcomes in patients. The advantages of remission derive from adding duration to absolute symptom score thresholds, and avoiding percentage change scores (a hitherto dubious benchmark).

CONCLUSIONS

Symptom rating scales which have been designed to diagnose patients, subdivide patients, define syndromes, track clinical change or evaluate drug efficacy do not lend themselves easily to the assessment of global outcome in schizophrenia. Simply totalling the number of symptoms without reference to the consequences of what is scored, is an empty exercise. Change must be relative to baseline conditions; there are also issues of redundancy, and a lack of concurrent validity with external outcome measures. The effort expended investigating the psychometric properties of scales such as the PANSS appears to have been matched by only limited advances in their utility beyond tracking change. It has yielded little of relevance to aetiology, treatment or prognosis.

Table 4 Proposed items for remission criteria with cross-scale correspondence and relationship to historical constructs of psychopathology dimensions and DSM–IV criteria for schizophrenia[1]

Dimension of psychopathology	DSM–IV criterion	SAPS and SANS items		PANSS items		BPRS items	
		Criterion	Global rating item no.	Criterion	Item no.	Criterion[2]	Item no.
Reality distortion	Delusions	Delusions (SAPS)	20	Delusions	P1	Grandiosity	8
				Unusual thought content	G9	Suspiciousness	11
						Unusual thought content	15
	Hallucinations	Hallucinations (SAPS)	7	Hallucinatory behaviour	P3	Hallucinatory behaviour	12
Disorganisation	Disorganised speech	Positive formal thought disorder (SAPS)	34	Conceptual disorganisation	P2	Conceptual disorganisation	4
	Grossly disorganised or catatonic behaviour	Bizarre behaviour (SAPS)	25	Mannerisms/ posturing	G5	Mannerisms/ posturing	7
Negative symptoms (psychomotor poverty)	Negative symptoms	Affective flattening (SANS)	7	Blunted affect	N1	Blunted affect	16
		Avolition–apathy (SANS)	17	Social withdrawal	N4	No clearly related symptom	
		Anhedonia – asociality (SANS)	22				
		Alogia (SANS)	13	Lack of spontaneity	N6	No clearly related symptom	

BPRS, Brief Psychiatric Rating Scale; PANSS, Positive and Negative Syndrome Scale; SANS, Scale for Assessment of Negative Symptoms; SAPS, Scale for Assessment of Positive Symptoms.
1. For symptomatic remission, maintenance over a 6-month period of simultaneous ratings of mild or less on all items is required.
2. Use of BPRS criteria may be complemented by use of the SANS criteria for evaluating overall remission.

These limitations have led to interest in another perspective on outcome, remission. This is based on accepted practice in medicine and other psychiatric disorders, such as affective disorders, and goes beyond rating scale scores alone. Its utility, however, remains to be seen. There are already indications that remission may be short lived in many patients (Dunayevich *et al*, 2006). Until recovery can be defined accurately in schizophrenia (Leucht & Lasser, 2006) symptom control, remission and quantified cognitive, personal and social functioning should be used together as measures of treatment outcome. This accepts that outcome has multiple facets, which vary in importance between patients. Symptom rating scales play an important role in overall appraisal of outcome, but should not dominate the picture, which still requires meaningful appraisals of cognition, personal and social functioning.

REFERENCES

Addington, D., Addington, J., Maticka-Tyndale, E., et al (1992) Reliability and validity of a depression rating scale for schizophrenics. *Schizophrenia Research*, **6**, 201–208.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorder* (4th edn). APA.

Andreasen, N. (1982) Negative symptoms in schizophrenia: definition and reliability. *Archives of General Psychiatry*, **39**, 784–788.

Andreasen, N. (1984a) *Scale for the Assessment of Negative Symptoms (SANS)*. Univeristy of Iowa.

Andreasen, N. (1984b) *Scale for the Assessment of Positive Symptoms (SAPS)*. University of Iowa.

Andreasen, N. & Olsen, S. (1982) Negative v positive schizophrenia: definition and validation. *Archives of General Psychiatry*, **39**, 789–794.

Andreasen, N., Carpenter, W., Kane, J., et al (2005) Remission in schizophrenia: proposed criteria and rationale for consensus. *American Journal of Psychiatry*, **162**, 441–449.

Andrews, G., Peters, L., & Teeson, M. (1994) *The Measurement of Consumer Outcome in Mental Health*. Australian Health Outcomes Collaboration.

Bell, M., Milstein, R., Beam-Goulet, J., et al (1992) The Positive and Negative Syndrome Scale and the Brief Psychiatric Rating Scale. Reliability, comparability and predictive validity. *Journal of Nervous and Mental Disorders*, **180**, 723–728.

Bell, M., Lysaker, P., Milstein, R., et al (1994) Concurrent validity of the cognitive component of schizophrenia: relationship of PANSS scores to neuropsychological assessment. *Psychiatry Research*, **54**, 51–58.

Beneke, M. & Rasmus, W. (1992) Clinical Global Impressions (ECDEU): some critical comments. *Pharmacopsychiatry*, **25**, 171–176.

Birsoz, S., Isik, E., Yargic, I., et al (2006) Remission in schizophrenia: a utopia or a goal? *Klinik Psikofarmakoloji Buelteni*, **16**, 128–132.

Bozikas, V., Kosmidis, M., Gamvrula, K., et al (2004) Clock Drawing Test in patients with schizophrenia. *Psychiatry Research*, **121**, 229–238.

Chadwick, P. & Birchwood, M. (1995) The omnipotence of voices. II: The Beliefs About Voices

Questionnaire (BAVQ). *British Journal of Psychiatry*, **166**, 773–776.

Collins, A., Remington, G., Coulter, K., et al (1996) Depression in schizophrenia: a comparison of three measures, *Schizophrenia Research*, **20**, 205–209.

Cramer, J., Rosenheck, R., Xu, W., et al (2001) Detecting improvement in quality of life and symptomatology in schizophrenia. *Schizophrenia Bulletin*, **27**, 227–234.

Crow, T. (1980) Positive and negative schizophrenic symptoms and the role of dopamine. *British Journal of Psychiatry*, **137**, 383–386.

d'Amato, T., Rochet, T., Dalery, J., et al (1992) Relationship between symptoms rated with the Positive and Negative Syndrome Scale and Brain Measures in schizophrenia. *Psychiatry Research*, **44**, 55–62.

Delmonte, P., Gabrielli, F., Gibertie, F., et al (1970) Diagnostic classification concepts among Italian psychiatrists. *Archivio di Psicologia Neurologia e Psichiatria*, **6**, 531–553.

Dunayevich, E., Sethuraman, G., Enerson, M., et al (2006) Characteristics of two alternative schizophrenia remission definitions: relationship to clinical and quality of life outcomes. *Schizophrenia Research*, **86**, 300–308.

Eckert, S., Diamond, P., Miller, A., et al (1996) A comparison of instrument sensitivity to negative symptom change. *Psychiatry Research*, **63**, 67–75.

Emsley, R., Oosthuizen, P., Joubert, A., et al (1999) Depressive and anxiety symptoms in patients with schizophrenia and schizophreniform disorder. *Journal of Clinical Psychiatry*, **60**, 747–751.

Emslie, G., Rush, A., Weinberg, W., et al (1997) A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Archives of General Psychiatry*, **54**, 1031–1037.

Engelsmann, F. & Formankova, M. (1967) Abbreviated Psychiatric Evaluation Scale (Overall-Gorham). *Ceskoslovenska Psychiatrie*, **63**, 297–306.

Engelsmann, F., Finar, O., Pichot, P., et al (1970) International comparison of diagnostic patterns. *Transcultural Psychiatric Research Review*, **7**, 130–137.

Fresan, A., De la Fuente-Sandoval, C., Loyzaga, C., et al (2005) A forced five-dimensional factor analysis and concurrent validity of the Positive and Negative Syndrome Scale in Mexican schizophrenic patients. *Schizophrenia Research*, **72**, 123–129.

Gharabawi, G., Rodriguez, S., Turkoz, I., et al (2005) Exploring the evidence for a relationship between symptomatic remission and functioning in patients with schizophrenia. *Neuropsychopharmacology*, **30** (suppl. 1), s1–s270.

Good, K., Rabinowitz, J., Whitehorn, D., et al (2004) The relationship of neuropsychological test performance with the PANSS in antipsychotic naive, first-episode psychosis patients. *Schizophrenia Research*, **68**, 11–19.

Guy, W. (1976) Clinical Global Impression. In *ECDEU Assessment Manual for Psychopharmacology* (revised), pp. 217–221. National Institute of Mental Health.

Hamera, E., Schneider, J., Potocky, M., et al (1996) Validity of self-administered symptom scales in clients with schizophrenia and schizoaffective disorders. *Schizophrenia Research*, **19**, 213–219.

Hamilton, S., Revicki, D., Genduso, L., et al (1998) Olanzapine versus placebo and haloperidol: quality of life and efficacy results of the North American double-blind trial. *Neuropsychopharmacology*, **18**, 41–49.

Haro, J., Kamath, S., Ochoa, S., et al (2003) The Clinical Global Impression-Schizophrenia scale: a simple instrument to measure the diversity of symptoms

present in schizophrenia. *Acta Psychiatrica Scandinavica*, **107** (suppl. 416), 16–23.

Hedlund, J. & Vieweg, B. (1980) The brief psychiatric rating scale (BPRS): a comprehensive review. *Journal of Operational Psychiatry*, **11**, 48–65.

Kane, J., Honigfeld, G., Singer, J., et al (1988) Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives of General Psychiatry*, **45**, 789–796.

Kay, S., Fiszbein, A., & Opler, L. (1987) The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, **13**, 262–276.

Kay, S., Opler, L., & Lindenmayer, J. (1988) Reliability and validity of the Positive and Negative Syndrome Scale for schizophrenia. *Psychiatry Research*, **23**, 99–110.

Kay, S., Opler, L., & Lindenmayer, J. (1989) The Positive and Negative Syndrome Scale (PANSS): rationale and standardisation. *British Journal of Psychiatry*, **155** (suppl. 7), 59–67.

Kissling, W., Heres, S., Lloyd, K., et al (2005) Direct transition to long-acting risperidone − analysis of long-term efficacy. *Journal of Psychopharmacology*, **19**, 15–21.

Krauss, H., Marwinski, K., Held, T., et al (1998) Reliability and validity of the premorbid adjustment scale (PAS) in a German sample of schizophrenic and schizoaffective patients. *European Archives of Psychiatry in Clinical Neuroscience*, **248**, 277–281.

Lasser, R., Bossie, C., Gharabawi, G., et al (2005) Remission in schizophrenia: results from a 1-year study of long-acting risperidone injection. *Schizophrenia Research*, **77**, 215–227.

Leucht, S. & Engel, R. (2006) The relative sensitivity of the Clinical Global Impressions Scale and the Brief Psychiatric Rating Scale in antipsychotic drug trials. *Neuropsychopharmacology*, **31**, 406–412.

Leucht, S. & Lasser, R. (2006) The concepts of remission and recovery in schizophrenia. *Pharmacopsychiatry*, **39**, 161–170.

Leucht, S., Kane, J., Kissling, W., et al (2005a) Clinical implications of Brief Psychiatric Rating Scale scores. *British Journal of Psychiatry*, **187**, 366–371.

Leucht, S., Kane, J., Kissling, W., et al (2005b) What does the PANSS mean? *Schizophrenia Research*, **79**, 231–238.

Leucht, S., Kane, J., Etschel, E., et al (2006) Linking the PANSS, BPRS, and CGI: clinical implications. *Neuropsychopharmacology*, 1–8.

Lindenmayer, J., Kay, S. & Friedman, C. (1986) Negative and positive schizophrenic syndromes after the acute phase: A prospective follow-up. *Comprehensive Psychiatry*, **27**, 276–286.

Lindenmayer, J., Bernstein-Hyman, R. & Grochowski, S. (1994) A new five factor model of schizophrenia. *Psychiatric Quarterly*, **65**, 299–322.

Lindenmayer, J., Brown, E., Baker, R., et al (2004) An excitement subscale of the Positive and Negative Syndrome Scale. *Schizophrenia Research*, **68**, 331–337.

Liu, S., Hwu, H. & Chen, W. (1997) Clinical sympton dimensions and deficits on the Continuous Performance Test in schizophrenia. *Schizophrenia Research*, **25**, 211–219.

Marshall, M., Lockwood, A., Bradley, C., et al (2000) Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *British Journal of Psychiatry*, **176**, 249–252.

Mass, R., Schoemig, T., Hitschfeld, K., et al (2000) Psychopathological syndromes of schizophrenia:

evaluation of the dimensional structure of the positive and negative syndrome scale. *Schizophrenia Bulletin*, **26**, 167–177.

McGorry, P., Goodwin, R. & Stuart, G. (1988) The development, use, and reliability of the brief psychiatric rating scale (nursing modification) − an assessment procedure for the nursing team in clinical and research settings. *Comprehensive Psychiatry*, **29**, 575–587.

Mortimer, A. & Spence, S. (2001) *The Management of Negative Symptons of Schizophrenia*. Science Press.

Murray, R. & Van Os, J. (1998) Predictors of outcome in schizophrenia. *Journal of Clinical Psychopharmacology*, **18**, 2S–4S.

Nasrallah, H. (2006) Improving patient outcomes in schizophrenia; achieving remission. *Journal of Psychopharmacology*, **20**, 57–61.

Norman, R., Malla, A., Cortese, L., et al (1996) A study of the interrelationship between and comparative interrater reliability of the SAPS, SANS and PANSS. *Schizophrenia Research*, **19**, 73–85.

Overall, J. & Gorham, D. (1962) The Brief Psychiatric Rating Scale. *Psychological Reports*, **10**, 799–812.

Overall, J. & Pfefferbaum, B. (1982) The Brief Psychiatric Rating Scale for Children. *Psychopharmacology Bulletin*, **18**, 10–16.

Overall, J. E. & Rhoades, H. M. (1982) Refinement of phenomenological classification in clinical psychopharmacology research. *Psychopharmacology*, **77**, 24–30.

Overall, J. & Beller, S. (1984) The Brief Psychiatric Rating Scale (BPRS) in geropsychiatric research: I. Factor structure on an inpatient unit. *Journal of Gerontology*, **39**, 187–193.

Overall, J. & Gorham, D. (1988) The Brief Psychiatric Rating Scale (BPRS): recent developments in ascertainment and scaling. *Psychopharmacology Bulletin*, **24**, 97–99.

Peralta, V. & Cuesta, M. (1994) Psychometric properties of the Positive and Negative Syndrome Scale (PANSS) in schizophrenia. *Psychiatry Research*, **53**, 31–40.

Perry, P., Sanger, T. & Beasley, C. (1997) Olanzapine plasma concentrations and clinical response in acutely ill schizophrenic patients. *Journal of Clinical Psychopharmacology*, **17**, 472–477.

Pogue-Geile, M. & Harrow, M. (1984) Negative and positive symptoms in schizophrenia and depression: A follow up. *Schizophrenia Bulletin*, **10**, 371–387.

Rabinowitz, J., Mehnert, A. & Eerdekens, M. (2006) To what extent do the PANSS and CGI-S overlap? *Journal of Clinical Psychopharmacology*, **26**, 303–307.

Ritsner, M., Blumenkrantz, H., Dubinsky, T., et al (2006) The detection of neurocognitive decline in schizophrenia using the Mindstreams Computerized Cognitive Test Battery. *Schizophrenia Research*, **82**, 39–49.

Rocca, P., Bellino, S., Calvarese, P., et al (2005) Depressive and negative symptoms in schizophrenia: different effects on clinical features. *Comprehensive Psychiatry*, **46**, 304–310.

Schuldberg, D., Quinlan, D. & Glazer, W. (1999) Positive and negative symptoms and adjustment in severely mentally ill outpatients. *Psychiatry Research*, **85**, 177–188.

Sethuraman, G., Taylor, C., Enerson, M., et al (2005) A retrospective comparison of cumulative time spent in remission during treatment with olanzapine or risperidone among patients with schizophrenia. *Schizophrenia Research*, **79**, 337–340.

s13

Singh, M. & Kay, S. (1975) A comparative study of haloperidol and chlorpromazine in terms of clinical effects and therapeutic reversal; with benztropine in schizophrenia: theoretical implications for potency differences among neuroleptics. *Psychopharmacologia*, **43**, 103–113.

Spearing, M., Post, R., Leverich, G., et al (1997) Modification of the Clinical Global Impression (CGI) scale for use in bipolar illness: the CGI-BP. *Psychiatry Research*, **73**, 159–171.

Steinert, T., Wolfle, M. & Gebhardt, R. (2000) Measurement of violence during in-patient treatment and association with psychopathology. *Acta Psychiatrica Scandinavica*, **102**, 107–112.

Tiihonen, J., Wahlbeck, K., Lonnqvist, J., et al (2006) Effectiveness of antipsychotic treatments in a nationwide cohort of patients in community care after first hospitalisation due to schizophrenia and schizoaffective disorder: observational follow-up study. *British Medical Journal*, **333**, 224.

Van den Oord, E., Rujescu, D., Robies, J., et al (2006) Factor structure and external validity of the PANSS revisited. *Schizophrenia Research*, **82**, 213–223.

van der Gaag, M., Cuijpers, A., Hoffman, T., et al (2006a) The five-factor model of the Positive and Negative Syndrome Scale I: confirmatory factor analysis fails to confirm 25 published five-factor solutions. *Schizophrenia Research*, **85**, 273–279.

van der Gaag, M., Hoffman, T., Remijsen, M., et al (2006b) The five-factor model of the Positive and Negative Syndrome Scale II: a ten-fold cross-validation of a revised model. *Schizophrenia Research*, **85**, 273–279.

Van Os, J., Fahy, T., Jones, P., et al (1996) Psychopathological syndromes in the functional psychoses: associations with course and outcome. *Psychological Medicine*, **26**, 161–176.

Van Os, J., Burns, T., Cavallero, R., et al (2006) Standardized remission criteria in schizophrenia. *Acta Psychiatrica Scandinavica*, **113**, 91–95.

Ventura, J., Neuchterlein, K. S. K., Gutkind, D., et al (2000) Symptom dimensions in recent-onset schizophrenia and mania: a principal components analysis of the 24-item Brief Psychiatric Rating Scale. *Psychiatry Research*, **97**, 129–135.

Welham, J., Stedman, T. & Clair, A. (1999) Choosing negative symptom instruments: issues of representation and redundancy, *Psychiatry Research*, **87**, 47–56.

White, L. (2005) Interpreting the PANSS: measures, factors and models. *Schizophrenia Research*, **79**, 349–351.

Wolthaus, J., Dingemans, P., Schene, A., et al (2000) Component structure of the Positive and Negative Syndrome Scale (PANSS) in patients with recent-onset schizophrenia and spectrum disorders. *Psychopharmacology*, **150**, 399–403.

ANN M. MORTIMER, BSc, MD, MMedSci, FRCPsych, Department of Psychiatry, University of Hull, Hertford Building, Cottingham Road, Hull HU6 7RX. Email: a.m.mortimer@hull.ac.uk