Outcome revealed by preference in schizophrenia (OPS): development of a new class of outcome measurements

BRUNO FALISSARD,¹ NADINE BAZIN,² MARIE-CHRISTINE HARDY-BAYLE²

1 Université Paris XI, AP-HP, INSERM U669, Paris, France

2 UVSQ, INSERM ERI 15, Centre Hospitalier de Versailles, Versailles, France

Abstract

The objective of this paper is to describe the development of a new type of outcome measurement, based on revealed preference, which can be used in serious chronic illnesses. Fifteen texts of about 200 words each were written by one of the authors on the day-to-day life of 15 schizophrenic patients. These 15 'slices of life' thus described were then ranked in terms of acceptability by a second group of 10 schizophrenic patients and by a group of 12 relatives of schizophrenic patients. From these rankings, six situations were selected so as to obtain evenly distributed positioning on an axis of acceptability. These six situations comprised the final instrument. In administration, the patients were first asked if the 'slices of life' that were described were acceptable or not, then if the 'slices of life' described were more or less acceptable than their own lives. Two scores were derived, one for an absolute level and the other for a relative level of the patient's satisfaction with his or her existence. Validation results were presented to a new sample of 229 schizophrenic patients. Internal consistency appeared good and the initial ranking of the six situations in terms of acceptability was confirmed. This study encourages the development of global outcome measures based on revealed preference in chronic serious illnesses. Copyright © 2006 John Wiley & Sons, Ltd.

Key words: schizophrenia, outcome assessment, health status indicator, rating scale, preference, validation study

Introduction

There is growing interest in developing outcome measurements in psychiatry, especially in serious chronic illnesses like schizophrenia. Measures currently in use, such as tests of cognition, and scales that measure clinical symptoms, functional status, satisfaction and quality of life, all attempt to capture different aspects of the broad concept of outcome. As such, they have significantly improved clinicians' ability to assess improvements in these domains on a quantitative basis. However shortcomings remain in the methodology currently available in this area.

Cognitive measurements, including cerebral functional imaging, have great scientific potential and are

able to measure small changes (Purdon et al., 2000). But even if cognitive dysfunction is increasingly recognized as a major contributor to the adaptive impairment seen in most patients with schizophrenia (McGurk et al., 2000), these measurements are sometimes criticized for presenting built-in biases (Loewenstein et al., 1989). Furthermore, they are designed to measure a relatively narrow aspect of outcome and cannot grasp many aspects of patients' existence, such as their place in society or their subjective wellbeing. For all these reasons, cognitive measurements can, at the moment, only be considered as promising surrogates.

Clinical signs and symptoms have been derived from the analysis and synthesis of clinical observations

conducted by several generations of psychiatrists; they have thus a certain level of historical or face validity. Moreover, because, typically, symptoms in schizophrenia are established from what patients express about their disease, they are likely to be correlated to the burden of disease, as indeed is the case in other diseases. Many good instruments such as the PANSS (Kay et al., 1989) are now available to measure this clinical domain. Measures based on signs and symptoms, however, may not necessarily be good proxies for assessing quality of life. Furthermore, physicians may often follow a 'logic of complaint': if there is no complaint from the patient or the family the physician may consider that the situation is acceptable for the patient. In schizophrenia, this may be an especially relevant concern because patients and especially those with prominent negative symptoms may be less likely to bring complaints to the physician's attention.

Functional status measures, which are in practice related to what the patient is able to perform (Cyr et al., 1994), are of particular interest because they are objective and fairly easy to assess. Furthermore, if a patient has a high level of functioning this is likely to be associated with a reasonably acceptable life. Thus functional status can be considered to be a relevant outcome measure. Functional status measurements, nevertheless, have several drawbacks. First, they most often lack responsiveness and it may take a good deal of time before a treatment leads to an improvement in, for instance, a patient's marital or professional status. Moreover, functional status and its relevance in terms of outcome is likely to depend on many variables, such as cultural considerations: different cultures may have different beliefs about what is considered to be optimal functioning in society (Sechrest and Flores, 1969).

An important objective of any treatment should be to improve patients' quality of life and to satisfy them. Some quality-of-life instruments have been designed in the field of schizophrenia (Lehman, 1983; Heinrichs et al., 1984; Auquier et al., 2003). However, these instruments have potential drawbacks.

First, there is a basic methodological concern. Quality-of-life instruments are obtained by a process of reduction: groups of patients are generally asked to detail what constitutes life satisfaction or quality of life for them (Kitzinger, 1995). This methodology presupposes that a human being is able to perform such an evaluation task, which may be especially questionable when the patient is impaired, or may generate

defensive reactions that interfere with introspective abilities.

Second, a normative conception of quality of life poses a philosophical problem. Some authors argue that, from an existentialist perspective, quality of life and satisfaction are contingent and can be assessed and interpreted only relative to a given moment, through the free expression of an enlightened patient (Lepège and Hunt, 1997). Thus any attempt to capture quality of life from predetermined answers to systematic items is likely to fail.

Finally, most of these instruments are multidimensional and thus lead to a profile of scores (Ware and Sherbourne, 1992). This is a real problem when assessing treatment efficacy where, for methodological reasons, there is a need for a single major endpoint.

All these criticisms justify the need for new developments in the area of outcome measurements in schizophrenia.

Here a new approach is proposed, based on a methodology of 'revealed preference', obtained from the ranking by the subject of six small written 'slices of life' of schizophrenic patients. When this instrument is administered, the patient is first asked if the individuals that are described in the 'slices of life' have an acceptable existence or not (on a four-point response scale) and then, in a second step, the patient is asked if the patients described have a more-or-less acceptable life than himself or herself (also on a four-point response scale).

Methods

Development of the instrument

The authors selected a sample of 15 schizophrenic patients from their clinical practice, aiming to provide a wide sample of outcomes. The concept of outcome was not explicitly defined, patients being selected on a qualitative basis in respect to how their life was going.

In a second step, the last author conducted a 1-hour interview with each of these 15 patients during which:

- the DSM IV (American Psychiatric Association, 1994) diagnosis of schizophrenia was confirmed;
- patients were asked to speak freely about their present lives.

After this interview, the last author summarized this 'slice of life' in about 170 words, plus 20 words providing

some elements of the patient's history. The age of the patient was arbitrarily fixed at 30. Two versions of the situations were written, one with a female patient and the other with a male patient. The six cases and the final scale are available on request or on the web at: <http://perso.orange.fr/bruno.falissard/Fichiers/OPSinstrument.doc>.

In a third step, three groups were recruited:

- a group of 10 schizophrenic patients recruited on the basis of a maximum heterogeneity in terms of age, gender and place of residence;
- a group of 12 relatives of schizophrenic patients, all members of the UNAFAM (French association of families and friends of patients with mental disorders);
- a group of 11 psychiatrists.

All members of these three groups were asked to rank the 15 situations according to their level of acceptability. To increase the reliability of the scores, two alternative ranking methods were used. First, each subject was asked, for each situation, to complete a Visual Analog Scale (VAS). The instruction was 'Place a mark on the line below according to how acceptable you feel the patient's situation is in the paragraph you have just read.' There were four labels above the 10cm horizontal line, from left to right: 'not at all acceptable', 'not really acceptable', 'fairly acceptable' and 'acceptable'. Second, the 15 situations were randomized and presented in groups of three. Each subject was then asked to rank each set of three situations by increasing level of acceptability. This two-step ranking procedure enabled

- a reliability coefficient to be estimated for each subject (see statistical section);
- a mean ranking to be estimated for each situation.

Subjects with reliability lower than 0.5 were removed.

In a fourth step, on the basis of these rankings, six situations were chosen, with the aim of obtaining evenly spaced intervals on an axis ranging from a very unacceptable situation to a very acceptable situation.

These six situations make up the instrument. In administration, an eight-page booklet is given to the patient in the presence of an investigator. The first page is an introduction that may be read to the patient, or

that the patient may read alone. The three following pages present the six situations (the gender of the patients in the described situation should be the same as the gender of the patient interviewed). They are each followed by the question: 'Tick one of the boxes below: Does Mr. H's life seem to you to be: "not acceptable at all", "not really acceptable", "fairly acceptable" or "acceptable"?' These four possible answers are associated to the numerical values 1, 2, 3 and 4. This first presentation of the situations is intended to familiarize patients with them and to assess if their answers are consistent with the rankings obtained during the development of the instrument. The situations are presented in an order of increasing level of acceptability; this was intended to help patients to position themselves among this succession of 'slices of life' in the second part of the instrument.

The fifth page is an introduction to the second part of the instrument:

In the following pages you will be reading 'slices of life' about the same 6 psychiatric patients as above. We would now like you to assess the lives of these different people in relation to your own life. In other words, we would like to know how far the life of each patient described is preferable or otherwise to your own life. It may sometimes be difficult to form a judgement, but please try to provide an answer each time.

Then the 6 situations are presented again, this time accompanied by the following questions: 'Tick one of the boxes below. Mr. H's life is: "much less acceptable than yours", "a little less acceptable than yours", "a little more acceptable than yours", "a lot more acceptable than yours"?' These four possible answers are associated with the numerical values 1, 2, 3 and 4. This second presentation of the situations is used to estimate the outcome.

Two outcome scores can be computed:

- the sum of the answers obtained for the second series of evaluations, where patients compare their existence to the six situations: this score relates to the absolute level of satisfaction of the patients with their existence;
- a second score can be computed by subtracting from the second score the summation of the answers obtained in the first series of evaluations.

This second score reflects the level of satisfaction of the patient relative to her/his expectations. This point is often considered as crucial in the field of satisfaction or quality-of-life measurements (Calman, 1984).

Elements of validation

The OPS is at present being used as a secondary end-point in a randomized controlled clinical trial comparing the safety and efficacy of two antipsychotics in two groups of 180 schizophrenic patients. Preliminary results obtained on the baseline evaluation of the first 229 patients are used here to provide certain elements of validation for the OPS.

Statistical analysis

Besides classic descriptive, bivariate and psychometrical analyses, this study required the use of less common techniques.

For the initial ranking of the situations, a reliability coefficient and a common estimator are estimated from the two rankings provided by each evaluator (VAS scores and 3 by 3 ranking). The common estimator is obtained from the VAS scores, permuted according to the 3 by 3 ranking procedure. For example if observations 1, 2, and 3 have been scored 23, 46 and 55 on the VAS and if their 3 by 3 ranking is 1 (first situation lowest), 3 (third situation intermediate) and 2 (second situation highest), the common scoring is 23 for situation 1, 55 for situation 2 and 46 for situation 3 (55 and 46 have been switched). This procedure has been specifically designed for the present experiment; it is justified by the fact that 3 by 3 rankings are assumed to be more reliable but only if limited to sets of three situations. The reliability coefficient is the Intraclass Correlation Coefficient of the common estimators and the VAS scores.

The selection of the 6 final situations among the original 15 was made using a multidimensional graphical representation (principal component analysis). In other words, since the first principal component summarizes the global level of acceptance of the 15 situations, the selection consists in taking situations equally distributed on this dimension.

All computations were performed with the 'R' software (R Development Core Team, 2004), with the 'base', 'psy' and 'boot' packages. All confidence intervals are bootstrapped with a confidence level of 95%.

Results

Construction of the instrument: choice of six situations out of 15

Among the 15 patients selected to produce the 'slices of life', there were 12 males and three females. Their ages ranged from 22 to 53.

One patient was not able to rank the situations and was removed from the study. The mean of the intraclass correlation coefficients of the two different rankings was equal to 0.47 in the patient group, 0.79 in the family group and 0.80 in the psychiatrist group. A boxplot is presented in Figure 1 to detail these results. Three patients and one psychiatrist had an ICC below 0.5 and were therefore removed from the rest of the analysis as provided for in the statistical analysis plan.

The internal consistency (with 95% confidence intervals) of patient rankings estimated with the alpha coefficient of Cronbach was 0.88 (0.46, 0.97). The internal consistency was 0.78 (0.11, 0.90) for the family members and 0.89 (0.51, 0.96) for the psychiatrists.

A graphical representation of the associations between rankings using a principal components analysis is given in Figure 2. Patient and family member evaluations are active while psychiatrist's evaluations are only illustrative: they do not contribute to the determination of principal components. On the right-hand part of the diagram the relationships between

Figure 1. Distribution of intra class correlation coefficients measured in three groups of evaluators: six schizophrenic patients, 12 family members of schizophrenic patients and 11 psychiatrists. These coefficients estimate, for each evaluator, the concordance between two methods of estimation of the level of acceptability of a series of 15 'slices of life' of schizophrenic patients.

Figure 2. Right: graphical representation, using a principal components analysis, of relationships between ratings of schizophrenic patients (Sub1, Sub2 . . .), schizophrenic patient family members (Fam1, Fam2 . . .) and psychiatrists (Psy1, Psy2 . . .). There is no obvious clustering of ratings according to the group to which the rater belongs.

Left: dual representation of the situations rated. Along the x-axis the situations are distributed according their level on the latent variable 'acceptability'.

rankings are presented. The x-axis explains a large amount of variance, 40%, which confirms the good consistency between patient and family member evaluations. In a representation of this type, this axis can be associated with a latent variable 'acceptability'. This means, for example, that family 1, patients 10 and 3 and psychiatrist 5 have an opinion concerning the level of acceptability of the situations that is close to the mean of the sample of 6 patients and 12 patient family members. This diagram also indicates, qualitatively, that there is no clear cluster of patients on one side, of families on another side and of psychiatrists elsewhere. In other words, it appears that there is no clear systematic difference in the assessments of the three groups of evaluators.

The left part of the diagram represents the 15 situations. It can be seen that situations C and L have been evaluated as being close. Along the x-axis the situations are distributed according to their level for the latent variable 'acceptability'. Situations that are circled were chosen to constitute the final instrument. The mean rankings of these situations are: 77 for H, 63 for K, 60 for Bo, 53 for Pi, 44 for T, and 30 for C. This confirms that these six situations are, on average, classified from the most acceptable to the least acceptable.

Elements of validation

As explained in the 'methods' section, the OPS is presently being used in a randomized controlled clinical trial comparing the efficacy and safety of two antipsychotics in schizophrenic patients (DSM IV).

Baseline data from 229 patients participating in this trial were available at the time of preparation of this manuscript; 225 provided complete answers enabling a preliminary analysis of the OPS. The internal consistencies are good for the six-item instrument: alpha is equal to 0.69 for the first series of six questions and alpha is equal to 0.72 for the second series.

The levels of acceptability of the six situations are 2.7 for H, 2.6 for K, 2.3 for Bo, 2.3 for Pi, 2.4 for T and 1.9 for C. Levels for Bo, Pi and T are not significantly different when a paired t-test is used to compare mean scores. Thus these 225 patients ranked the six situations in the same way as the patients and patient family members in the first part of the study.

Discussion

This paper presents a new approach in the assessment of outcome of schizophrenic patients. Many methodological aspects are not innovative and may be related to Thurstone's method of equal appearing intervals

(Thurstone and Chave, 1929) and utility revealed by preference (Drummond et al., 1997). Indeed, preferences are regularly used in the development of quality of life measurements: the EuroQol is an example (EuroQol Group, 1990), and a feasibility study of the paradigm of utility revealed by preference has been published in the field of schizophrenia (Voruganti et al., 2000).

The present approach is new in the sense that the outcome is not reduced to a list of items. It is obtained from the assessment by the patient under evaluation of 'slices of life' that come from real schizophrenic patients and that have been ranked by patients and schizophrenic patient family members. Patients under evaluation can position themselves over a range of situations and the score obtained from this process gives an indirect quantification of the level of acceptability of the patients' existence, according to their own impressions. The assessment can be relative: acceptability as compared to the life of other patients. The assessment can also be absolute because a first series of questions asks the patient if the 'slices of life' are acceptable or not in themselves: the combination of both series of answers is an estimate of an absolute assessment.

The preliminary validation results presented are encouraging, indicating acceptable to good internal consistency, and the ranking of the 'slices of life' selected for the scale in terms of acceptability has been replicated in the validation sample of patients. However, caution in the interpretation of these results is needed, as observations were presented in predefined order of decreasing level of acceptability and thus an ordering effect cannot be ruled out. It should be noted that the ranking does not need to be strictly constant across patients; what is more important is that patients under evaluation have the opportunity of expressing themselves in comparison with a wide range of other patients' existences.

Many other limitations may be pointed out. The patients described in the 'slices of life' have been chosen by clinicians, the observations themselves have been written by a psychiatrist, so that a medical viewpoint bias may be present in the descriptions. However, this point was discussed with the groups of patients and patient family members and no comment of this sort was reported. Of course this is not a conclusive argument. Another limitation could be poor responsiveness: this question will be answered when the results from the ongoing clinical trial are available.

Finally, if an English version of the instrument is given here for strictly informative purposes, this is not to say that this instrument has been validated for English-speaking countries. Indeed, it is possible that situations considered as acceptable or otherwise may differ from culture to culture, so that it is recommended to develop a specific instrument in each particular setting.

Acknowledgements

This study has been supported by an unconditional grant from Eli Lilly France. The authors want to thank the patients and the UNFAM (French association of families and friends of patients with mental disorders) for their help in the ranking of the 'slices of life'.

References

- American Psychiatric Association. DSM-IV. Diagnostic and Statistical Manual of Mental Disorders. 4 edn. Washington DC: American Psychiatric Association, 1994.
- Auquier P, Simeoni MC, Sapin C, Reine G, Aghababian V, Cramer J, Lancon C. Development and validation of a patient-based health-related quality of life questionnaire in schizophrenia: the S-QoL. *Schizophr Res* 2003; 63(1–2): 137–49.
- Calman KC. Quality of life in cancer patients – an hypothesis. *J Med Ethics* 1984; 10(3): 124–7.
- Cyr M, Toupin J, Lesage AD, Valiquette CA. Assessment of independent living skills for psychotic patients. Further validity and reliability. *J Nerv Ment Dis* 1994; 182(2): 91–7.
- Drummond MF, O'Brien BJ, Stoddard GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes* Oxford: Oxford University Press, 1997.
- EuroQol Group. EuroQol – a new facility for the measurement of health related quality of life. *Health Pol* 1990(16): 199–208.
- Heinrichs DW, Hanlon TE, Carpenter WT, Jr. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. *Schizophr Bull* 1984; 10(3): 388–98.
- Kahneman D, Tversky A. Choices, values, and frames. *Am Psychol* 1984; 39: 341–50.
- Kay SR, Opler LA, Lindenmayer JP. The Positive and Negative Syndrome Scale (PANSS): rationale and standardisation. *Br J Psychiatry Suppl* 1989(7): 59–67.
- Kitzinger J. Qualitative research. Introducing focus groups. *BMJ* 1995; 311(7000): 299–302.
- Lehman AF. The well-being of chronic mental patients. *Arch Gen Psychiatr* 1983; 40(4): 369–73.
- Lepège A, Hunt S. The problem of quality of life in medicine. *JAMA* 1997; 278(1): 47–50.
- Loewenstein DA, Amigo E, Duara R, Guterman A, Hurwitz D, Berkowitz N, Wilkie F, Weinberg G, Black B,

- Gittelman B, Eisdorfer C. A new scale for the assessment of functional status in Alzheimer's disease and related disorders. *J Gerontol* 1989; 44(4): 114–21.
- McGurk SR, Moriarty PJ, Harvey PD, Parrella M, White L, Davis KL. The longitudinal relationship of clinical symptoms, cognitive functioning, and adaptive life in geriatric schizophrenia. *Schizophr Res* 2000; 42(1): 47–55.
- Purdon SE, Jones BD, Stip E, Labelle A, Addington D, David SR, Breier A, Tollefson GD. Neuropsychological change in early phase schizophrenia during 12 months of treatment with olanzapine, risperidone, or haloperidol. The Canadian Collaborative Group for research in schizophrenia. *Arch Gen Psychiatry* 2000; 57(3): 249–58.
- R[computer program]. Version 2.1.1. <http://cran.r-project.org>; 2004.
- Sechrest L, Flores L. Homosexuality in the Philippines and the United States: the handwriting on the wall. *J Soc Psychol* 1969; 79: 3–12.
- Thurstone LL, Chave EJ. *The Measurement of Attitudes*. Chicago: University of Chicago Press, 1929.
- Voruganti LN, Awad AG, Oyewumi LK, Cortese L, Zirul S, Dhawan R. Assessing health utilities in schizophrenia. A feasibility study. *Pharmacoeconomics* 2000; 17(3): 273–86.
- Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30(6): 473–83.
- Correspondence:* Bruno Falissard, INSERM U669, Maison de Solenn, 97 Boulevard du Port-Royal, 75679 Paris Cedex 14, France.
 Telephone (+33) 1 58412848
 Fax (+33) 1 4559 3418
 Email: falissard_b@wanadoo.fr