International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 18(4): 265–278 (2009)
Published online 25 August 2009 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/mpr.287

An empirical method to identify patterns in the course of psychotic episodes of people with schizophrenia

LORENZO BURTI,[1] MARIANGELA MAZZI,[1] MAARTEN KOETER,[2] AART SCHENE,[2] HEDDA HELM,[3] BERND PUSCHNER,[4] JONATHAN BINDMAN,[5] MORVEN LEESE,[5] GRAHAM THORNICROFT[5] & MICHELE TANSELLA[1]

1 Section of Psychiatry and Clinical Psychology, Department of Medicine and Public Health, University of Verona, Verona, Italy
2 Department of Psychiatry, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands
3 Department of Psychiatry, University of Leipzig, Leipzig, Germany
4 Mental Health Research Unit, Department of Psychiatry, University of Ulm, Guenzburg, Germany
5 Institute of Psychiatry, King's College London, London, UK

Key words
schizophrenia, clinical course, factor analysis, cluster analysis

Correspondence
Lorenzo Burti, Section of Psychiatry and Clinical Psychology, Department of Medicine and Public Health, University of Verona, Policlinico G. B. Rossi, P.le L. A. Scuro 10, 37134 Verona, Italy.
Telephone (+39) 045-812-7441
Fax (+39) 045-812-4997
Email: lorenzo.burti@univr.it

Received 18 January 2008;
revised 2 April 2008;
accepted 13 June 2008

Abstract

Objective: This paper illustrates the process of constructing, selecting and applying simple measures in order to empirically derive patterns of course of psychotic episodes in schizophrenia.
Method: Data were collected with a composite instrument constructed for a multi-centre, follow-up randomized controlled trial of adherence therapy for people with schizophrenia. The instrument included a retrospective weekly assessment of psychotic/non-psychotic status, which was used to derive the measures, and the DSM-IV course specifiers.
Results: The measures discriminated well between different course patterns and identified homogeneous clusters of subjects which correlated with the groups derived from the DSM-IV course specifiers.
Conclusions: The new measures provide an empirical basis to identify specific patterns of course and to differentiate patients according to pre-defined criteria. They can be used in follow-up studies as measures of outcome, to investigate correlations between variables and to identify potential predictors of outcome. *Copyright © 2009 John Wiley & Sons, Ltd.*

Introduction

Longitudinal studies have overturned the historical beliefs on schizophrenia, originally regarded as a chronically deteriorating illness and showed instead that heterogeneous courses were common. These findings were compatible with the alternative view that different courses may belong to different sub-types of schizophrenia, or even to different illnesses (Bleuler, 1972; Ciompi and Müller, 1976; Huber *et al.*, 1980; Harding, 1988;

McGlashan, 1988). Various numbers of discrete patterns have been described by these and other authors, mostly based on clinical judgment or simple operational definitions, to differentiate between episodic or continuous course, like in the Life Chart Schedule (LCS) where, unfortunately, symptom duration was not rated because this proved unfeasible (Susser *et al.*, 2000).

An ordinary way of describing course patterns is part of DSM-IV [Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)]: the clinical course specifiers (American Psychiatric Association, 2000). They include five profiles differentiating continuous and episodic course, with the latter split into single- and multiple-episode types, with inter-episodic either full or partial remission, but without further definition of the duration of episodes. The sixth profile is reserved for other or unspecified patterns, to be indicated by the rater himself, thus providing useful but unclassified information. These profiles do not have a time scale: they just provide a qualitative appraisal of the overall course without any anchoring time points. While they may be of interest in the long run or over a life-time course, they are of little help in outcome studies of a definite follow-up duration. They also do not give the opportunity to work with statistics, as needed in an outcome study.

This paper illustrates an empirical method to identify course profiles embodying time references and providing information on course trends. This method is based on the data derived from a retrospective weekly assessment of patient condition (psychotic/non-psychotic). The resulting patterns are then correlated with those employing the DSM-IV course specifiers, i.e. patterns based on clinical judgement. Clearly, the course of psychotic illness is multidimensional, but some aspects of the course can be described in terms of 'on' and 'off' phenomena, such as positive psychotic symptoms. This paper focuses on a way to describe the longitudinal course of such 'on/off' phenomena in psychotic illness, i.e. the course of psychotic episodes within psychotic illness. The method does not apply to other, less discrete phenomena such as negative symptoms, disability and depression, other aspects of the course of psychotic illness.

This study uses the data collected with a composite instrument to describe the course of psychotic episodes, devised for the QUATRO study, a 12-month, single-blind, multi-centre randomized controlled trial (RCT) of the effectiveness of adherence therapy compared to a supportive control intervention for people with an acute episode of schizophrenia in four European cities (London, Leipzig, Verona, Amsterdam – Santander was also involved in part of the study) (Gray *et al.*, 2006).

Methods

Subjects and settings

For the QUATRO study a total of 409 ICD-10 subjects [International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10)] with a clinical diagnosis of schizophrenia were recruited in the four sites; diagnoses were confirmed using the Item Group Checklist (IGC) of the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (Wing, 1992). In brief: out of 1218 people screened, 917 were eligible to participate, 366 (39.9%) refused to participate and 142 (15.5%) could not be randomized for other reasons. A total of 409 (44.6%) people were randomized. Baseline and follow-up data for the core outcome measures were collected for 349 (85.3%) participants: 184 (90%) in the health education group and 165 (81%) in the adherence therapy group, a difference in follow-up rate that was statistically significant ($p = 0.01$). People who dropped out of the trial tended to have had more in-patient days ($p = 0.022$), but in other respects were similar to those who completed the interviews, and the drop-outs were similar in the two arms (Gray *et al.*, 2006).

The focus of the study was upon the improvement of the quality of life and the reduction in disability of people in the European Union (EU) region who suffer from schizophrenia. The following domains were investigated with currently used standardized instruments: demographics, psychopathology, disability, patient and family quality of life, medication, side-effects and adherence, and costs. Instruments were administered both at baseline and at the 12-month follow-up.

Instrument

As to the course of psychotic episodes, an instrument including DSM-IV course specifiers, the 'Clinical Course Assessment for the QUATRO study' (CCA-EU) was devised, tested, validated and administered at follow-up. The instrument had to be expressly created, since none of the kind existed when the QUATRO study was developed. The CCA-EU is a semi-structured instrument to guide the raters to identify psychotic episodes, the circumstances of their occurrence and hospitalizations in the course of *positive* psychotic symptomatology during the selected period: in the case of the QUATRO study it was administered at the follow-up interview and covered the previous 12 months. Raters were experienced clinicians who were trained in the use of the instrument during the pilot phase of the study. Interrater reliability resulted in the high range: agreement 0.8–1.00; Kappa 0.44–1.00.

The raters were requested to use the best source of information available for an individual patient, i.e. either the patient him/herself, the patient's keyworker and/or casenotes. It was the responsibility of the raters to decide in a particular case which information or combination of information was the best. The following guidelines were provided: (i) when the patient was able to give adequate answers also on the other instruments, patient information was considered adequate, i.e. valid and reliable; (ii) when a keyworker had been in regular contact with a patient for the 12 month period and he/she had made case notes during this period, the keyworker was considered as the most reliable source of information available. However, sometimes keyworker´s information might not provide reliable information. For example: the keyworker had not had regular appointments with the patient; the patient was referred to another keyworker or service; the patient had lost contact with the service; (iii) Casenotes: available for the 12-month period and of good enough quality. At the end of the questionnaire the raters were asked to mention the source(s) of information used and to rate the quality of information in terms of the reliability of the findings (CCA-EU items 7 to 10: yes/no questions).

The instrument provides an operational definition of psychotic episode, or relapse. A patient is defined as having a psychotic episode when *one or more* of the following selected items of the Brief Psychiatric Rating Scale (BPRS) (Ventura *et al*., 1993) have a score of four to seven (moderate to extremely severe): hallucinations, unusual thought content or delusions, suspiciousness, conceptual disorganization, mannerism, bizarre behaviour; the duration of these symptoms is minimally one week, and two psychotic episodes need to be distinguished by at least four weeks without psychosis. The instrument includes three parts:

(i) Twelve initial items (CCA-EU 1–12) in the form of yes/no questions record whether the patient is in a psychotic episode according to definition at baseline, during, and at the end of the follow-up; periods spent in a psychotic episode and an appraisal of the reliability of the source of information;

(ii) In the second part of the instrument, a week-by-week evaluation of the patient being either psychotic or non-psychotic according to definition (a BPRS score of four to seven on these symptoms) is retrospectively performed (items CCA-EU 13/1-48: in fact each month was conventionally considered to include four weeks). Hospitalizations, if present, are also recorded in the same way (items CCA-EU 14/1-48: Figure 1).

(iii) The third part consists of the DSM-IV course specifiers. Rating is anchored to the definition for psychotic episodes and to DSM-IV specifiers for pattern of course (items CCA-EU 15/1-6). Original definitions are supplemented with pictorials to facilitate rater's selection (Figure 2).

For the purposes of this study only parts (ii) and (iii) were used.

Data analysis

Identification of different patterns of course of psychotic episodes

Leffondré *et al*. (2004) suggest basic statistical measures that are able to discriminate between different longitudinal patterns of change, like stable–unstable, or increasing–decreasing. Several measures have to be initially formulated, since no single one may encompass the various aspects of change to differentiate patients. A process of selection then follows, in order to avoid redundancy: the number of measures is reduced by means of factor analysis and, finally, the resulting measures are used to group patients with similar patterns using cluster analysis.

Initial construction of measures

Measures have to detect patients' specific characteristics like persistence of a stable clinical condition, for example, or the occurrence and duration of exacerbations. In our instrument, the sequence of weeks with or without positive psychotic symptoms according to definition, in the follow-up period, make it possible to use various types of

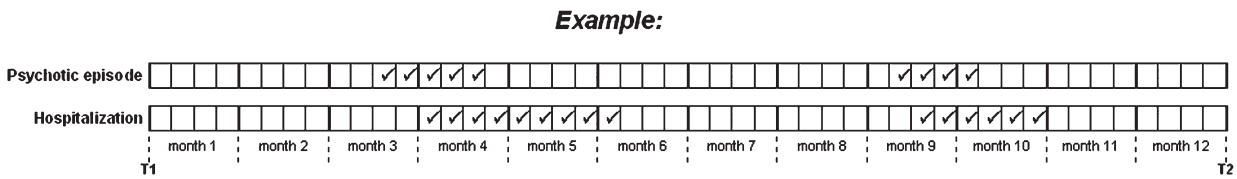

Figure 1 Example of the recording of weeks passed in a psychotic episode and/or admissions to hospital using the form provided in the instrument.

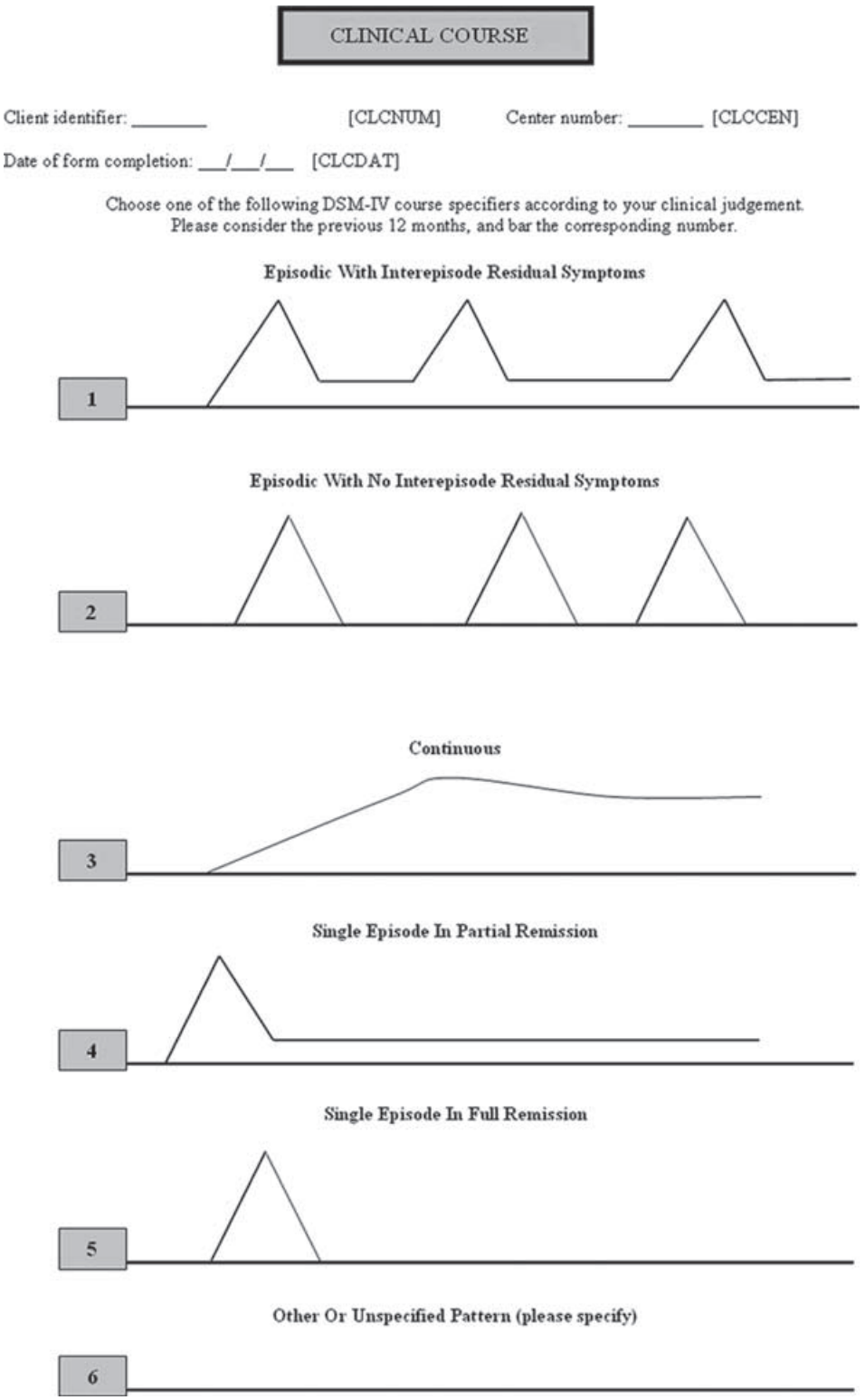

Figure 2 The Clinical Course Assessment for the QUATRO study (CCA-EU), part three: The DSM-IV course specifiers with pictorials.

measures *appropriate to binary data* like those listed in Table 1:

- Measures based on the number of weeks with positive psychotic symptoms during follow-up over the whole period, per quarter, in the first quarter compared to the rest, in the last quarter compared to the rest; in the third and fourth quarters compared to the first and second quarters; difference between first and last week. Comparisons across quarters and weeks provide information about change and its direction – improvement or deterioration – over time in the follow-up period. Corresponding measures of variability are also calculated in order to uncover peculiar aspects of course patterns.
- Measures of change: a simple count of the number of contiguous positive (zero to one) and negative (one to zero, were one means: 'presence of positive psychotic symptoms', zero means 'absence of positive psychotic symptoms'), or of any kind, weekly changes and related derived measures. We used '*Gower's binary similarity coefficient*' (detailed in the Appendix) instead of the odds ratio, because of our skewed distribution (Gower, 1985). It varies between zero and one; it has a value of one when there is no change, zero when all consecutive observations vary. This measure provides fundamental information whether a patient has a continuous or an episodic course and differentiates between courses like those of DSM-IV specifiers No. 3 (continuous course) and Nos 1–2 and 4–5 (episodic course). However, it does not provide any temporal information on timing and temporal sequence of episodes, if any. Other measures of this type may be calculated on contiguous periods longer than one week e.g. on two-, three- and four-week periods.

Table 1 Descriptive analyses of 14 relevant measures on 335 subjects

Variable	Mean	Standard deviation (SD)	Minimum	Maximum	Coefficient of variation (CV)
Measures based on the number of weeks with positive psychotic symptoms during follow-up					
Weeks with positive psychotic symptoms in the whole period	16.38	19.74	0	48	1.20
Variation coefficient of weeks with positive psychotic symptoms in the whole period	1126.64	1385.94	0	4749.74	1.23
Difference of number of weeks with positive psychotic symptoms between first and last semester	−0.07	4.69	−20	20	−67.00
Variance of weeks with positive psychotic symptoms in first quarter	0.56	0.93	0	3	1.66
Variance of weeks with positive psychotic symptoms in second quarter	0.23	0.69	0	3	3.0
Variance of weeks with positive psychotic symptoms in third quarter	0.34	0.85	0	3	2.5
Variance of weeks with positive psychotic symptoms in last quarter	0.47	1.01	0	3	2.15
Measures of change					
Number of changes absence-to-presence of symptoms	0.43	0.77	0	6	1.79
Number of changes presence-to-absence of symptoms	0.58	0.80	0	6	1.38
Gower in whole period	0.89	0.15	0	1	0.17
Gower in first semester	0.92	0.13	0	1	0.15
Gower in last semester	0.93	0.15	0	1	0.16
Short-term fluctuations from symptomatic to non-symptomatic and vice-versa					
Number of short periods (two weeks) with symptoms	0.04	0.23	0	2	5.75
Number of short periods (two weeks) without symptoms	0.01	0.08	0	1	8.00

- Duration of periods of stable clinical conditions: longest period with zero (absence of positive psychotic symptoms), longest period with one (presence of positive psychotic symptoms). Such measure provides temporal information on duration of either stability or episode(s) of relapse.
- Short-term fluctuations from symptomatic to non-symptomatic and vice-versa: number of two-week episodic sequences with or without positive psychotic symptoms, i.e. having the form of 0110 or 1001 (one-week sequences like 010 and 101 were too rare in our sample to be used). This measure may confirm whether the course is either continuous or episodic and provides information on the temporal sequence of episodes.
- Other measures may be constructed in order to further refine the ability of the method to identify a broader set of patterns. While the construction of other measures goes beyond the purposes of the instrument designed for the QUATRO study, nonetheless the method allows the addition of more of them in the future, to pursue further research aims.

Selection of measures

Since the resulting measures may be redundant and overlapping, a subsequent *selection phase* is necessary to eventually supply a more contained set. To this purpose, inter-measure correlations were calculated to identify redundant measures ($r > 0.9$) among which the clinically relevant ones were selected (see example in results). A factor analysis was then performed on the residual measures; only factors with an eigenvalue greater than 1.0 were retained. The varimax rotation was employed and, according to the resulting rotated factor loading matrix, for each factor, the measure with the highest loading was selected.

Classification of patterns

These measures were then used to catalogue each subject's pattern into relatively homogeneous groups with cluster analysis using the method of *k*-means in order to assign each subject to one group (Everitt *et al.*, 2001). Prior to performing the cluster analysis, the measures, expressed in different units, were standardized with the *Z*-score transformation. Stata software package V. 9.1 was used for data analysis. This whole procedure followed that of Leffondré *et al.* (2004). Sensitivity analysis was then performed in order to test the robustness of the results (Kettenring, 2006).

Results

The number of subjects with valid data in the two relevant sections of the instrument (sections two and three) was 335. Of these, 195 (58%) were male; 234 (70%) single; 175 (52%) with secondary education; 52 (16%) in paid employment; 134 (40%) unemployed; 107 (32%) retired. Average age was 41 years [standard deviation (SD) = 11.5; range = 18–69]. Mean duration of illness 13 years (SD = 9; range = 0–47). Baseline clinical assessment: average Global Assessment of Functioning (GAF) Scale total score was 50 (SD = 14; range = 9–85); average BPRS total score was 45 (SD = 13; range = 24–99).

As to the sources of information, in 60% of subjects two sources were used (in 52% one source was the keyworker); in 38% one source only was used (in 27% the source was either the keyworker or the casenotes and only in 11% the patient alone). As to the quality of information, in 97% of subjects the interviewers rated the quality as either very or generally reliable.

Weekly evaluation of subjects

The means of weeks with positive psychotic symptoms and/or hospitalizations per month were higher at month one, possibly in relation to participants' inclusion criteria, i.e. 'clinical instability'. In fact they were 1.70 (SD = 1.8) and 0.75 (SD = 1.4), respectively. In the following months (2–12) they remained substantially stable, in the range 1.27–1.50 (SD = 1.8–1.9) and 0.31–0.45 (SD = 1.1–1.2), respectively.

DSM-IV course specifiers

The number of subjects for each of the six groups identified using the DSM-IV course specifiers are reported in the rightmost column of Table 2. Note that the highest numbers of subjects belong to group III (*continuous*: $N = 95$) and VI (*other or unspecified pattern*: $N = 84$). This includes a majority of subjects with a flat profile, indicating a condition of non-psychotic most of the time, besides a minority with irregular patterns.

Construction of measures to identify course patterns, using the weekly evaluations of patients

Following Leffondré's suggestions, we created measures with the aim of designing empirical profiles of the course of psychotic episodes. In our analysis 14 clinically relevant measures to discriminate between patients with different patterns of course were selected out of 55 measures

Table 2 Percentage distribution of subjects classified according to DSM-IV group specifiers by group derived from cluster-analysis

	Groups derived from cluster-analysis					
DSM-IV course specifier	Subjects psychotic most of the time	Subjects non-psychotic most of the time	Subjects psychotic in first semester	Subjects psychotic in last semester	Subjects with fluctuating course	Number of subjects
I – Episodic, inter-episodic residual symptoms	1.9	1.9	5.7	32.1	58.5	53
II – Episodic, no inter-episodic residual symptoms.	0.0	15.4	15.4	30.8	38.5	13
III – Continuous	80.0	5.3	5.3	4.2	5.3	95
IV – Single episode in partial remission	1.7	28.3	56.7	11.7	1.7	60
V – Single episode in full remission	0.0	10.0	50.0	36.7	3.3	30
VI – Other or unspecified pattern	10.7	82.1	2.4	4.7	0.0	84

applicable to our dichotomous data originally identified. This initial selection was performed by the means of a bivariate table of correlation; in the case of overlapping measures, the more clinically specific were selected. An example follows. Three measures of change were initially selected: number of changes presence-to-absence of symptoms, indicating improvement, number of changes absence-to-presence of symptoms, indicating deterioration; and number of changes, total, indicating change in either direction, i.e. instability. The observed correlation between the measure indicating improvement (number of changes presence-to-absence of symptoms) and instability (number of total changes) showed redundant information (Pearson's $r = 0.94$). Also the correlation between deterioration (number of changes absence-to-presence of symptoms) and instability was high (Pearson's $r = 0.93$). From a statistical point of view, the choice to discard a variable among these is neutral, because when two variables are known, the last one can be calculated (instability is the sum of the other two variables). From a clinical point of view, instead, the direction of change, either improvement or deterioration, is more informative than instability; instability, then, seemed to be the best candidate for removal.

The 14 resulting measures are presented in Table 1 with their mean, SD, range and coefficient of variation (CV). A single measure may identify a patient with a specific profile, as in the case of the measure 'Weeks with positive psychotic symptoms in the whole period' where the value 48 identifies the profile of subject 1192 (always psychotic), and zero that of subject 1254 (always *non*-psychotic). In all other cases a combination of more measures is necessary to identify a specific pattern. As an example, the profiles of six participants' different patterns of course are provided (Figure 3).

For this example, we selected subjects showing course profiles possibly relevant for the study of reference, i.e. the QUATRO study. The corresponding most characterizing

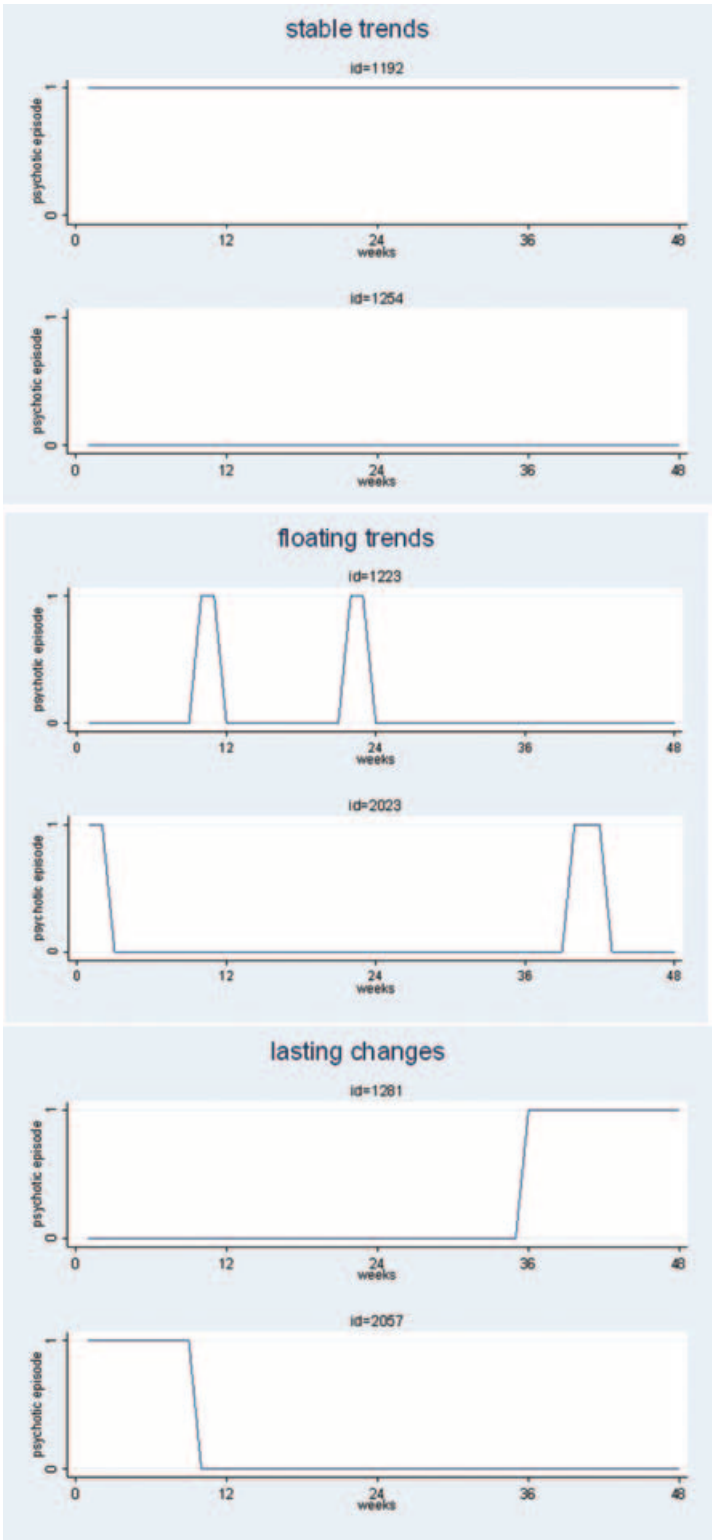

Figure 3 Profiles of six participants with different patterns of course provided as an example.

Int. J. Methods Psychiatr. Res. 18(4): 265–278 (2009). DOI: 10.1002/mpr

Table 3 Factor analysis – rotated factor loadings

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Gower in first semester	0.94			
Variance of weeks with positive psychotic symptoms in first quarter	0.83			
Gower in the whole period	0.70	0.62		
Number of changes presence-to-absence of symptoms	0.67		0.54	
Gower in last semester		0.93		
Difference of number of weeks with positive psychotic symptoms between first and last semester	0.52	−0.71		
Number of short periods (two weeks) with symptoms			0.93	
Number of changes absence-to-presence of symptoms		0.64	0.61	
Weeks with positive psychotic symptoms in the whole period				0.79
Variance of weeks with positive psychotic symptoms in second quarter	0.47			0.53

Note: The factor loadings smaller than 0.4 (in absolute value) are not displayed.

measures among the 14 previously selected (Table 1) are also reported:

- stable clinical profiles with continuous course, either always psychotic (subject 1192), or never psychotic (subject 1254). The characterizing measure is *number of weeks with positive psychotic symptoms*, which amounts to 48 and 0, respectively;
- psychotic episodes (two) in the first semester (subject 1223) that lasted four weeks total and implied two changes absence-to-presence (measure: *Number of changes absence-to-presence of symptoms*);
- psychotic status at baseline and a psychotic episode in the second semester (subject 2023). Two measures contribute to describe this subject, namely: *number of weeks with positive psychotic symptoms* (5), *difference of number of weeks with positive psychotic symptoms between first and last semester* (−1);
- steadily non-psychotic at baseline and during most of the follow-up period; ensuing psychotic episode continuing through follow-up evaluation (deteriorating pattern: subject 1281); relevant measures are: *weeks with positive psychotic symptoms in the whole period* (13), *difference of number of weeks with positive psychotic symptoms between first and last semester* (−13);
- steadily psychotic at baseline and for the initial part of the follow-up period; non-psychotic for most of the follow-up period and at follow-up (improving pattern: subject 2057); relevant measures are: *weeks with positive psychotic symptoms in the whole period* (9), *difference of number of weeks with positive psychotic symptoms between first and last semester* (9).

Selection of measures with factor analysis

Table 3 illustrates the rotated factor loading matrix, which indicates the correlation between each variable and the four maintained factors.

- Factor 1 relates to the variability of patterns in general, and especially in the first semester, in that it is characterized by variances: it discriminates a stable pattern from a variable one;
- Factor 2 relates to the variability of patterns in the second semester;
- Factor 3 relates to short-term fluctuations;
- Factor 4 relates to the number of weeks with positive psychotic symptoms.

These four factors explain 83% of variance. The Kaiser–Meyer–Olkin measure of sampling adequacy is 0.77.

At this point we moved on to identifying the most informative variables in the pool included in Table 3. We identified four measures by choosing the prevalent one in each factor: Gower in the first semester, Gower in the last semester: i.e. whether a patient had a continuous or episodic course in the first or in the last semester, respectively; number of short periods (two weeks) with positive psychotic symptoms; number of weeks with positive psychotic symptoms in the whole period.

Cluster analysis

Using the four variables just identified, we created relatively homogeneous groups of subjects with respect to these variables that best discriminated the subjects

according to their different course patterns. We resolved to use five clusters basically for clinical reasons, since we were interested in identifying stable subjects, either psychotic or non-psychotic; subjects who either improved or deteriorated and subjects with a fluctuating course. A clustergram was used to confirm the number of clusters.

The results of cluster analysis are presented in Table 4, which reports the descriptive statistics, i.e. mean, SD, median for each cluster and for the total sample. The numbers of weeks with positive psychotic symptoms per quarter are also reported on the right side of the table for reference. In order to test the robustness of our cluster analysis, this was replicated 10 times while modifying the random number seed of the original randomization. Total concordance (10/10) was obtained in 59.4% of our sample, while good concordance (eight replications out of 10) was obtained in another 26% of the sample. In this case, the difference is attributable to cluster 1 which, on two occasions, aggregated with cluster 2. So, 85% of our sample aggregated in a stable way; in particular, stability resulted greatest for clusters 2 and 3.

Comparison between the DSM-IV course specifiers and the results of the cluster analysis

In order to compare the classification of course patterns using DSM-IV course specifiers with those derived from cluster analysis, Table 2 presents the percentage of subjects of each DSM-IV group sharing the participation in each of the five groups derived from cluster analysis.

The comparison shows a fair correspondence among the course profiles derived by using two different methods. Note that the subjects belonging to DSM-IV course specifiers I and II, are distributed among three groups derived from cluster analysis, namely subjects psychotic in first semester, subjects psychotic in last semester, subjects with fluctuating course. In fact, while DSM-IV course specifiers do not provide time anchoring, groups derived from cluster analysis do. The same is true for DSM-IV course specifiers IV and V. However, the method described in this paper does not discriminate between subjects who recover completely between episodes and those who do not: an asset of DSM-IV course specifiers. However, this is just a limit deriving from the crude psychotic/non-psychotic distinction employed in this particular study. The method allows the use of a rating scale of psychotic symptoms. Note that a comparison between two methods of rating the same phenomenon shows possible inconsistencies. For example, only 80% of subjects belonging to DSM-IV course specifier III (continuous course) belongs to the cluster *subjects psychotic most of the time*; the remaining 20% is scattered among the other groups. The reasons for such inconsistencies between different parts of the instrument will be discussed in a future paper dealing with the whole instrument.

Discussion

This study uses the data collected with two out of the three parts of an instrument (the CCA-EU) constructed and employed to assess the 12-month course of people with schizophrenia participating in a multi-centre RCT of adherence therapy (QUATRO study). It aims to assess the week-by-week evaluation of the patient being either psychotic or non-psychotic, and the DSM-IV course specifiers. The validation of the *parent* CCA-EU instrument itself will be presented somewhere else. Fourteen measures describing the course of psychotic episodes have been selected out of the 55 initially created. The goal was one of eliminating redundancies. This was initially performed on the basis of correlation. In the case of measures with a high correlation indicating they are measuring the same dimension, clinical criteria were used for selection.

The measures had to discriminate between different course patterns and correctly identify the specific patterns of individual patients for both screening purposes of patients and the assignment of patients to groups with similar course. Patterns of interest in our study were those of either improvement or relapse, stability or fluctuation. Today, patterns over time may be assessed with more sophisticated methods than those suggested by Leffondré *et al.* (2004), e.g. the latent growth methods (LGM: Skrondal and Rabe-Hesketh, 2004). This statistical technique has a number of strengths and can be applied to any data of repeated measures but it needs a *confirmatory* perspective, as in the case when a definite set of hypotheses is tested to support a theory. Although the LGM has several merits over traditional techniques for the analysis of change, in this study we have preferred an *exploratory* strategy of analysis, following the suggestions of Leffondré, rather than a theory-driven one.

Factor analysis

The most informative variables resulting from factor analysis were the following:

- variables describing *change* in general and, specifically, *in the 1st semester*, as *Gower in the 1st semester*, the *variance of weeks with positive psychotic symptoms in the first quarter*, and then the *Gower on the entire follow-up period* (in Factor 1);

Table 4 Descriptive statistics by cluster group (mean, SD, median, respectively, for each cluster and for the total sample) and number of weeks with positive psychotic symptoms per quarter for reference

Cluster	Gower in first semester	Gower in last semester	Weeks with positive psychotic symptoms in the whole period	Number of short periods with positive psychotic symptoms	Weeks with positive psychotic symptoms in 1st quarter	Weeks with positive psychotic symptoms in 2nd quarter	Weeks with positive psychotic symptoms in 3rd quarter	Weeks with positive psychotic symptoms in 4th quarter
1 – Subjects psychotic most of the time ($n = 87$; 26%)	1	1	47.6	0	12	12	12	11.6
	0	0	0.98	0	0	0	0	0.98
	1	1	48	0	12	12	12	12
2 – Subjects non-psychotic most of the time ($n = 97$; 29%)	0.99	0.98	0.22	0	0.18	0.01	0.02	0.01
	0.11	0.15	0.48	0	0.38	0.10	0.14	0.10
	1	1	0	0	0	0	0	0
3 – Subjects who are psychotic in the first semester, then recover ($n = 61$; 18%)	0.78	0.99	6.11	0	3.98	1.34	0.41	0.39
	0.09	0.04	7.17	0	3.21	2.94	2.16	2.15
	0.78	1	3	0	2	0	0	0
4 – Subjects who are non-psychotic in the first semester, then relapse ($n = 47$; 14%)	1	0.76	7.68	0.15	0.57	0.53	2.45	4.13
	0	0.14	7.63	0.42	2.45	2.45	3.15	3.67
	1	0.82	5	0	0	0	1	4
5 – Subjects with a fluctuating course ($n = 43$; 13%)	0.75	0.77	13.23	0.19	2.44	3.37	3.58	3.84
	0.11	0.13	7.5	0.45	2.64	4.12	3.82	3.32
	0.76	0.81	12	0	2	2	3	4
Total ($n = 335$)	0.92	0.93	16.4	0.01	4.29	3.87	4	4.22
	0.13	0.15	19.7	0.23	5.14	5.37	5.29	5.23
	1	1	6	0	2	0	0	0

- variables describing *change* in the *second* semester, thus highlighting the course of the second part of the follow-up period with regard to variability. They include *Gower in the 2nd semester* and the *difference of the number of weeks with positive psychotic symptoms between the first and the last semester* (in Factor 2);
- variable dealing with rapid fluctuations, which best relates to relapses. This is the *number of short periods (two weeks) with positive psychotic symptoms* (in Factor 3);
- variable *weeks with positive psychotic symptoms in the whole period,* which is related to persistent severity (in Factor 4).

These variables are consistent with the patterns of interest of our follow-up study on people with schizophrenia, the QUATRO study (improvement or relapse; stability or fluctuation): they are appropriate to describe our population which includes subjects who improved, subjects who relapsed; subjects with an illness characterized by long-periods of stability, and subjects with a fluctuating course. The most informative variables were then used to create clusters of subjects with similar courses. We propose this method for future follow-up studies on courses of psychotic illnesses; however appropriate change variables applicable to the specific study and setting shall be chosen. In addition, maybe less frequent, but *prospective* ratings of symptomatology are advisable.

Cluster analysis

The first group ($N = 87$) includes those subjects with a continuous course with positive psychotic symptoms and the second one ($N = 97$) those with a continuous course described as non-psychotic according to definition. These two groups together include more than half of the QUATRO subjects. Incidentally, the overall size of such, clinically stable, sub-population may be considered among the potential reasons for the lack of effect of the intervention tested in the study (Gray *et al*., 2006). The third group ($N = 61$) includes subjects who were initially psychotic and then recovered over the follow-up period, while the fourth ($N = 47$) includes subjects who started as non-psychotic and relapsed over time. The fifth group is composed of subjects with a fluctuating course of illness. Each cluster includes *relatively* homogeneous courses but the model allows also the identification of definite individual patterns, as shown in the examples. Thus, the proposed method provides an empirical basis for the construction and identification of course patterns like the DSM-IV course specifiers, as shown in the comparison of the two methods of classification. This method is also able to additionally construct more precise patterns of interest that may be of use for a range of schizophrenia research areas, ranging from genetic studies that want to investigate the modulating effects of gene variations on the longitudinal course of schizophrenia, to imaging and cognition studies that want to investigate the association between the course of psychotic episodes with structural and functional brain changes. In the case of use in a psychosocial or pharmacological trial, the method provides time anchoring of possible changes in the patterns: an important aspect when improvement or relapse and deterioration over time are of interest. In addition, the method can rely on a limited number of variables and, when opportunely paired with appropriate instruments of assessment, a limited number of observations. Finally, while the method was used over a follow-up period only 12-month long to identify psychotic episodes, it can be used over longer time periods to characterize also the medium- to long-term course of illness, by using more measures (i.e. not only positive symptoms) at wider intervals. This is a critical point to decide the type of interventions and to optimize the resources of mental health services (Di Michele *et al*., 2007), especially today when the therapeutic goal for schizophrenia has become more ambitious, i.e. the achievement of a relative degree of social and relational remission (Ruggeri and Tansella, 2008). However, this last point implies the caveat that one cannot generalize the analysis and use the instrument for follow-up studies longer than one year without previously checking the findings with a separate study spanning a more extended period of time.

Limitations

A major limitation of this study is the retrospective assessment of weekly *positive* psychotic symptoms, which is only one aspect of the course of schizophrenia, over an entire year duration. The distinction between psychotic and non-psychotic is based on the difference between a BPRS rating of three (mild severity) and four (moderate severity), a criterion of questionable precision. However, such difference is what clinicians are familiar with: either a symptom is mild and not of real clinical significance, or it is moderate, and therefore of clinical significance. In addition, this assessment is combined with the criterion of duration (at least one week) and the interrater reliability resulted in the high range.

Furthermore, keyworker, casenotes, interview with the subjects, not the family, were available as sources of information in the majority of, but not in all, cases; the analysis was neither preceded by a study on the validation

of the instrument, nor by an appraisal of reliability of the best source of information used by raters. However, the main aim of this paper is one of presenting a methodology using the QUATRO study just as an example, without pretending to analyse and reliably describe the clinical course of QUATRO subjects providing valid data.

Another limitation has to do with cluster analysis, which depends on a number of subjective choices: different choices may lead to different results. We chose a partitional cluster analysis with *k*-means because of the size of our sample and because our variables are continuous. In addition, one has to bear in mind that cluster analysis is a process based on a heuristic criterion and moves on from an arbitrary starting point in measuring distances. This implies the incomplete reproducibility of results: launching it again and again may lead to different results and the groups thus obtained are not completely overlapping. However, in our case, a ten-fold replication of the cluster analysis provided evidence for satisfactory robustness. Another limit has to do with choosing the optimal number of resulting groups. In our case we decided to limit the number of clusters to five on the basis of clinical reasons, a graphic criterion and the aim of obtaining groups of reasonable size.

Concluding Remarks

This study aimed to applying an empirical method to the description of course in schizophrenia and to differentiate patients according to selected criteria. This method makes it easy to screen also large amounts of data because it does not require clinical assessors' time-consuming and subjective evaluation of individual life charts. It can be used in follow-up studies of given duration, because it incorporates the time dimension and provides information regarding trends toward either improvement or deterioration, and also to investigate correlations between variables and to identify potential predictors of outcome.

Acknowledgements

The QUATRO study is a multi-centre collaboration between the Health Services Research Department, Institute of Psychiatry, King's College London; the Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology; University of Verona, Italy; the Department of Psychiatry, Leipzig University, and the Department of Psychiatry II, Ulm University, Germany; the Department of Psychiatry, Academic Medical Centre, University of Amsterdam, The Netherlands and Santander, Spain. The study was funded by a grant from the Quality of Life and Management of Living Resources Programme of the European Union (QLG4-CT-2001-01734). The views expressed in this publication are those of the authors and not necessarily those of the funders. We also wish to acknowledge the contributions of the patients, carers, and staff who have taken part in this study. We would like to acknowledge the contributions to this study of the following colleagues: Amsterdam site (Aart Schene, Annemarie Fouwels, Martijn Kikkert, Maarten Koeter, Karin Meijer); Leipzig/Ulm site (Thomas Becker, Matthias Angermeyer, Anja Born, Anne Gießler, Hedda Helm, Bernd Puschner); London site (Jonathan Bindman, Jayne Camara, Anthony David, Richard Gray, Martin Knapp, Morven Leese, Paul McCrone, Mauricio Moreno, Anita Patel, Debbie Robson, Graham Thornicroft, Ian White); Santander site (José Luis Vázquez-Barquero, Sara Herrera Castañedo, Verónica Chinchón, Andrés Herrán, Javier Vázquez-Bourgon); Verona site (Michele Tansella, Francesco Amaddeo, Corrado Barbui, Lorenzo Burti, Daniela Celani, Doriana Cristofalo, Claudia Goss, Antonio Lasalvia, Giovanna Marrella, Mariangela Mazzi, Michela Nosè, Mirella Ruggeri, Marta Solfa).

Declaration of interest statement

All authors declare that there are no conflicts of interest.

Policy and ethics

The work described in this article has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans; the protocol has been approved by individual ethical-committees at each site.

References

American Psychiatric Association. (2000) *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*, American Psychiatric Association.

Bleuler M. (1972) *Die schizophrenen Geistesstörungen im Lichte langjähriger Kranken – und Familiengeschichten*, Thieme.

Ciompi L., Müller C. (1976) *Lebensweg und Alter der Schizophrenen. Eine katamnestiche Langzeistudie bis ins Senium*, Springer.

Di Michele V., Bolino F., Mazza M., Roncone R., Casacchia M. (2007) Relapsing versus non relapsing course of schizophrenia: a cohort study in a community based mental health service. *Epidemiologia e Psichiatria Sociale*, **16**, 50–58.

Everitt B.S., Landau S., Leese M. (2001) Cluster analysis, 4th ed., Arnold.

Gower G.C. (1985) Measures of similarity, dissimilarity, and distance. In: *Encyclopaedia of Statistical Sciences* (eds Kotz S., Johnson N.L., Read N.), pp. 397–405, Wiley.

Gray R., Leese M., Bindman J., Becker T., Burti L., David A., Gournay K., Kikkert M., Koeter M., Puschner B., Schene A., Thornicroft G., Tansella M. (2006) Adherence therapy

for people with schizophrenia. European multicentre randomised controlled trial. *British Journal of Psychiatry*, **189**, 508–514, DOI: 10.1192/bjp.bp.105.019489.

Harding C.M. (1988) Course types in schizophrenia: an analysis of European and American studies. *Schizophrenia Bulletin*, **14**, 633–643.

Huber G., Gross G., Schüttler T., Linz M. (1980) Longitudinal studies of schizophrenic patients. *Schizophrenia Bulletin*, **6**, 592–605.

Kettenring J.R. (2006) The practice of cluster analysis. Journal of Classification, **23**, 3-30, DOI: 10.1007/s00357-006-0002-6.

Leffondré K., Abrahamowicz M., Regeasse A., Hawker G.A., Badley E.M., McCusker J., Belzile E. (2004) Statistical measures were proposed for identifying longitudinal patterns of change in quantitative health indicators. *Journal of Clinical Epidemiology*, **57**, 1049–1062, DOI: 10.1016/j.jclinepi.2004.02.012.

McGlashan T.H. (1988) A selective review of recent North American long-term followup studies of schizophrenia. *Schizophrenia Bulletin*, **14**, 515–542.

Ruggeri M., Tansella M. (2008) Improving the treatment of schizophrenia in real world mental health services. *Epidemiologia e Psichiatria Sociale*, **17**, 249–253.

Skrondal A., Rabe-Hesketh S. (2004) *Generalized latent variable modeling*, Chapman & Hall/CRC.

Susser E., Finnerty M., Mojtabai R., Yale S., Conover S., Goetz R., Amador X. (2000) Reliability of the life chart schedule for assessment of the long-term course of schizophrenia. *Schizophrenia Research*, **42**, 67–77.

Ventura J., Green M., Schaner A., Liberman R. (1993) Training and quality assurance with the Brief Psychiatric Rating Scale. The 'drift busters'. *International Journal of Methods in Psychiatric Research*, **3**, 221–244.

Wing J.K. (1992) *Schedules for Clinical Assessment in Neuropsychiatry*, World Health Organization.

Appendix

Binary similarity coefficient

$$\text{Binary similarity coefficient} = \frac{ad}{\sqrt{(a+b)(a+c)(d+b)(d+c)}}$$

where a is the number of concordant observations on presence of positive psychotic symptoms in contiguous weeks; d is the number of concordant observations on absence of positive psychotic symptoms in contiguous weeks; $(a + b)$ = number of times when the sequence starts with a week with positive psychotic symptoms; $(c + d)$ = number of times when the sequence starts with a week without positive psychotic symptoms; $(a + c)$ = number of times when the sequence ends with a week with positive psychotic symptoms; $(b + d)$ = number of times when the sequence ends with a week without positive psychotic symptoms.

Int. J. Methods Psychiatr. Res. 18(4): 265–278 (2009). DOI: 10.1002/mpr