Psychiatry Research 228 (2015) 551–559

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Investigating the long-term course of schizophrenia by sequence analysis

Wolfram an der Heiden*, Heinz Häfner

Schizophrenia Research Group, Central Institute of Mental Health, Mannheim Faculty of Medicine/University of Heidelberg, J5, D-68159 Mannheim, Germany

ARTICLE INFO

Article history:
Received 12 June 2014
Received in revised form
7 January 2015
Accepted 26 April 2015
Available online 19 June 2015

Keywords:
Schizophrenia
Long-term course
Outcome
Course types
Sequence analysis
Cluster analysis
Schizophrenia syndromes

ABSTRACT

In the present study we set out to explore the long-term clinical course of schizophrenia in a holistic manner by adopting sequence analysis. Our aim was to identify course types of illness by means of cluster analysis. The study was based on course and outcome data for 107 patients followed up over 134 months after first admission in the ABC Schizophrenia Study. Focusing on the main syndromes (positive, negative, depressive and unspecific symptoms) and their combinations we looked for similarities in individual illness courses using the 'optimal matching' method. A cluster analysis performed on the resulting similarity matrix yielded two main groups (an 'improving' and a 'chronic' group), which comprised a total of six different types of illness course. The course types differed in both quantitative (frequency of syndromes and syndrome combinations) and qualitative terms (clinical presentation, sequence of syndromes). Cluster membership was only rarely, but clearly associated with socio-demographic characteristics, treatment data and other illness variables.

1. Introduction

Today, illness course continues to be one of the leading topics in schizophrenia research. While Kraepelin used to regard a chronic course as a pathognomonic feature of the disorder, we now believe that, at a given point in time after onset, its presentation can be characterized by any of the variants ranging from an inconspicuous state to the most severe impairment. The heterogeneity of the disorder is reflected not only in cross-sectional psychopathology, but also in different types of illness course and outcome.

A key characteristic of schizophrenia is its episodic course. The episodes are defined primarily, but not exclusively, by positive symptoms. Negative symptoms and the amount of social impairment patients present are also factors taken into account. The pattern of how episodes and phenomenologically inconspicuous intervals alternate appears to be so characteristic of the disorder that attempts have been made to divide the heterogeneous group of schizophrenias into more homogeneous subgroups (Bleuler, 1972; Huber et al., 1979; Watt et al., 1983). In the end, however, all these efforts to differentiate various types of course have failed, more or less (Häfner and an der Heiden, 2003). The main reason is that the studies included in the analyses have been too heterogeneous regarding (a) the diagnostic criteria applied, (b) the constructs used for describing illness course, (c) the patient populations assessed and (d) the study designs.

The quality of studies for analysing and documenting the long-term illness course essentially depends on how data collection is designed. Empirical conclusions on the long-term course of schizophrenia are mostly drawn from data collected by a 'real-time- or a ‚catch-up' prospective design (cf. Robins, 1979); after entry in the study, the probands will be examined at pre-defined points in time to assess the variables of interest. To ensure the reliability and validity of the data collected commonly accepted assessment tools are often used. However, most psychopathological interviews yield reliable data only for the past 2–4 weeks preceding the interview. The time samples thus obtained provide the basis for describing the illness course.

A premise of that approach is that the cross-sectionally measured characteristics of the underlying concept are representative of the individual illness courses. However, as cross-sectional assessments tend to be scheduled taking practical considerations rather than how the illness evolves into account and given the time and staff needed for carrying them out, they are usually kept to a minimum, which means that such an approach cannot be regarded as truly valid. According to our meta-analysis of 72 long-term longitudinal schizophrenia studies covering 10 years or more merely an average of 1.7 follow-ups were conducted (an der Heiden, 1996).

Since the number of follow-ups cannot be easily stepped up for economic reasons, the main solution to the problem is to choose instruments specially developed for assessing psychopathology over lengthy periods of time. Examples for such instruments are:

* Corresponding author. Tel.: +49 621 1703 2952; fax: +49 621 1703 2955.
E-mail addresses: wolfram.ad_heiden@zi-mannheim.de, w.adheiden@gmail.com, heinz.haefner@zi-mannheim.de (W. an der Heiden).

http://dx.doi.org/10.1016/j.psychres.2015.04.056
0165-1781/

Schedule for Affective Disorders and Schizophrenia – Lifetime Version (SADS-L; Andreasen et al., 1981); Renard Diagnostic Interview (Helzer et al., 1981), Past History Schedule in conjunction with the Present State Examination (PHS/PSE; McGuffin et al., 1986), Diagnostic Interview Schedule (DIS; Robins et al., 1982), Composite International Diagnostic Interview (CIDI; Robins et al., 1988), Instrument for the Retrospective Assessment of the Onset and Course of Schizophrenia and Other Psychoses (IRAOS; Häfner et al., 1999, 2003). However, it is noteworthy that with the exception of the IRAOS none of the instruments listed were developed specifically for the purpose of reconstructing the illness course so as to enable longitudinal analyses.

A further key factor besides the method of collecting data on the relevant construct variables is the technique of statistical data processing. The method chosen should fit the longitudinal nature of the illness course and be appropriate for analysing the succession of events/episodes over long periods of time. Time-series analysis or the 'life table' method focuses on individual data points or transitions between them. Our aim in the present paper is to analyse complete data series. The following is a description of the first attempt ever to employ in epidemiological schizophrenia research a method originally developed in genetics for sequencing DNA strands: the sequence analysis.

2. Methods

2.1. Analysis of the long-term course of schizophrenia in the ABC Schizophrenia Study

The Age-Beginning-Course (ABC) Schizophrenia Study was launched in 1987, and the initial assessment was completed in 1989 (Häfner et al., 2013). The final follow-up assessment was done in the period from 1999 to 2002. The study sample comprised all German-speaking persons aged 12–59 years who were consecutively admitted to inpatient treatment for the first time with a diagnosis of schizophrenia at one of the 10 psychiatric hospitals serving a semi-urban, semi-rural population of 1.5 million. The sample has been described in detail elsewhere (e.g. Häfner et al., 1990).

Of the 276 patients initially recruited in the first-admission sample, 232 (84%) had not previously experienced any psychotic symptoms for more than 14 days (=first-episode sample) (Fig. 1). A subsample of around 56% ($n=130$) of the first-admission sample, of whom 115 patients were first-episode cases, were not only assessed at first admission, but also followed up at 6 months, 1, 2, 3 and 5 years later.

12.3 years (range: 11.2–14.6 years) after initial assessment 107 patients completed the long-term follow-up interview. 24 patients had died, 11 patients could not be traced, and 90 patients refused to cooperate. A comparison of the follow-up sample with the original sample did not reveal any significant differences with respect to important demographic and illness variables.

The main instrument of the ABC study, the 'Interview for the Retrospective Assessment of the Onset of Schizophrenia' (IRAOS), was used both at initial assessment and follow-up. Originally designed for the assessment of beginning schizophrenia (Häfner et al., 1990), the IRAOS has been revised and can now be used for assessing the course of all types of psychotic disorder (Häfner et al., 1999, 2003).

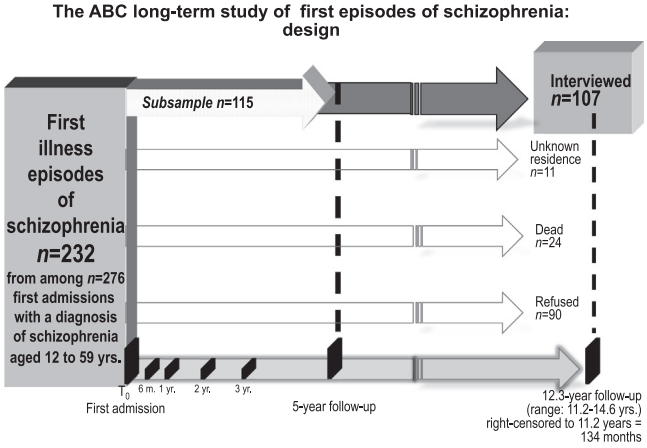

Fig. 1. ABC long-term study of first episodes of schizophrenia: design.

The following areas are covered in the IRAOS:

- Signs and symptoms – including unspecific and prodromal ones and deviant behaviour.
- Psychiatric impairment/social disability and functional impairment.
- Social indicators.
- Help-seeking behaviour (contacts with counselling and health services) and therapeutic interventions.

A key property of the instrument is that it permits a retrospective assessment and description of these indicators for an indefinite period of time according to features such as time of onset, type of course and duration. Since for each of the 126 symptoms time of onset and duration is determined, time series extending from first admission to long-term follow-up can be created.

2.2. Sequence analysis

In epidemiological research sequence analysis represents a heuristic, exploratory technique of comparing individual historical data. Originally developed in genetics for the purpose of ascertaining the degree of similarity between two DNA strands (Kruskal, 1983), sequence analysis is now used primarily in social sciences to study the sequence of events. The technique has been applied to issues such as transition from school to the labour market (e.g Anyadike-Danes and McVicar, 2010; Brzinsky-Fay, 2007; Scherer, 2001), professional career trajectories (e.g. Abbott and Hrycak, 1990; Pollock et al., 2002) or life course development (Billari, 2001; Martin et al., 2008).

A characteristic of sequence analysis is that entire series of data rather than individual data points are analysed.[1] Sequences or categorical time series are series of elements in a temporal or spatial order with each element or time point depicting a physical object or a specific state. For example, in career research such an element could be the state of being unemployed at a certain point in time, in genetic research a DNA base pair at a particular locus of a chromosome. The most important characteristic is the specific order of the elements, which cannot be changed. The purpose of sequence analysis is to compare two sequences in order to determine their similarity.

The technique most widely used for comparing sequences is the 'optimal matching' (OM) method. OM defines similarity as the minimum number of operations needed for transforming one sequence into another. The main operations involved are insertion (=i), deletion (=d) and substitution (=s).

To give an example: let us assume that there are two patients with schizophrenia, and we are comparing the following two sequences in their illness course over a period of 10 months following first admission to inpatient treatment:

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Pb1	pos	neg	pos	neg	pos	neg	neg	unsp	inco	inco
Pb2	pos	pos	neg	pos	neg	neg	unsp	unsp	unsp	inco
	0	1	1	1	1	0	1	0	1	0

pos: positive symptoms
neg: negative symptoms
unsp: unspecific symptoms
inco: inconspicuous

A direct comparison of their illness state across the months shows that there is dissimilarity in six of the 10 months (Hamming distance=6).

When the OM method is applied, the following picture emerges:

Pb1	pos	neg	pos	neg	pos	neg	neg	unsp		inco	inco
Pb2	pos		pos	neg	pos	neg	neg	unsp	unsp	unsp	inco
	0	i	0	0	0	0	0	0	i	s	0

i: gap insertion
s: substitution

[1] The following description of sequence analysis is mainly based on Brzinsky-Fay and Kohler (2010). For a more detailed discussion see Kruskal (1983).