

A longitudinal evaluation of two-year outcome in a community-based mental health service using graphical chain models

The South-Verona Outcome Project 9

MIRELLA RUGGERI,¹ GIULIA BISOFFI,¹ ANTONIO LASALVIA,¹ FRANCESCO AMADDEO,¹ CHIARA BONETTO,¹ ANNIBALE BIGGERI²

¹Department of Medicine and Public Health, Section of Psychiatry, University of Verona, Italy

²Department of Statistics, University of Florence, Florence Italy

ABSTRACT *The 2-year outcome of 178 patients attending a community-based mental health service was assessed from a multidimensional perspective. The study investigated: (1) the effect of disease-related characteristics (such as diagnosis and illness duration) and of a series of outcome variables measured at baseline (global functioning, psychopathology, social disability, quality of life and satisfaction with services) on total costs of care over 2 years; and (2) the effect of costs of care and outcome variables measured at baseline on the corresponding outcome variables at 2 years.*

To gain insight into the multivariate longitudinal dependencies among variables, we used graphical Gaussian chain models, a new multivariate method that analyses the relationship between continuous variables taking into account the effect of antecedent and intervening variables, to reveal not only direct but also indirect correlations.

Outcome variables showed the tendency to segregate, both at baseline and follow-up, into two distinct groups: a clinician-rated dimension (given by global functioning, social disability and psychopathology) and a patient-rated dimension (given by service satisfaction and subjective quality of life). Higher costs at 2 years were predicted by higher psychopathology at baseline, diagnosis of psychosis and longer duration of illness. Baseline values for each variable were the main predictors of the corresponding values at two years. Improvement in satisfaction with life at follow-up was experienced in those subjects with a lower functioning at baseline. This study throws some light on the complex relationships between clinical, social and economic variables affecting the medium-term outcome of mental health care.

Key words: community mental health services, outcome evaluation, longitudinal studies, multivariate analyses

Introduction

The effect of mental health care should be evaluated both longitudinally and multidimensionally – using a prospective follow-up design and exploring a broad range of outcome variables such as psychopathology, social disability, quality of life, service satisfaction and service utilization (Attkisson et al., 1992; Ruggeri and Tansella, 1995; Thornicroft and Tansella, 1999). Understanding the interrelationships among all these variables over time is crucial to the development of hypotheses on the relative importance of each of these indicators within routine clinical practice.

However, the concurrent investigation of complex sets of outcome variables may overshadow their mutual interaction unless appropriate data analysis techniques are used, because standard statistical techniques are unsatisfactory for this purpose. To deal with the complexity of interaction among variables, and to gain insight into the phenomena investigated, integrated methods for data representation and analysis are needed. Recent advances in the statistical literature have led to the identification of graphical modelling as a valuable tool for addressing research hypotheses within multivariate data (Whittaker, 1990; Lauritzen, 1996; Cox and Wermuth, 1996;

Edwards, 2000). The underlying idea is that 'any kind of relationship should be analysed as a conditional relationship' (Kreiner, 1996). For mental health services research this means that we cannot fully understand the determinants of outcome without taking into account the pattern of relationships among its multiple dimensions. *Graphical chain models* are a form of graphical model that is suitable for use in longitudinal studies, because they allow the analysis of complex sets of variables, some of which are predictors, some responses, and some intermediate between them.

Graphical chain models were first applied in mental health research in a 6-month follow-up study conducted by our group on a sample of patients attending the South Verona Community Mental Health Service (CMHS) and proved to be a useful strategy to analyse complex set of process and outcome data (Ruggeri et al., 1998a). As far as we know, few applications of this statistical methodology were attempted in mental health service research. The major limitation of this study, however, was the short period of investigation.

In order to overcome this limitation, we expanded the time span of investigation and followed up the sample referred to in the early study for 2 years. The present study assessed the 2-year multidimensional outcome of a cohort of patients attending the CMHS by using graphical chain models. Specifically, we aimed to investigate:

- the effect of both clinical patient's characteristics (e.g. illness duration, diagnosis) and of a series of outcome variables measured at baseline (e.g. global functioning, psychopathology, social disability, subjective quality of life and satisfaction with services) on total costs of care over 2 years; and
- the effect of costs and of above mentioned outcome variables measured at baseline on the corresponding outcome variables at 2 years.

Method

Setting

Data were collected in south Verona, a mainly urban area in the north-east of Italy, which is relatively affluent and predominantly middle class, with a low migration rate. The total population is about 75,000 inhabitants. The CMHS is the main mental health

service providing care to the south Verona residents. It includes a number of comprehensive and well-integrated services, and provides inpatient care, day care, rehabilitation, outpatient care and home visits, as well as a 24-hour emergency service and residential facilities for long-term patients. The same staff (excluding nurses) work in the hospital and community facilities and this 'single staff module' improves continuity of care over facilities and over time when compared with separate staff in community and hospital services. A Psychiatric Case Register (PCR), the South Verona Psychiatric Case Register, covers the same geographic area and collects demographic, diagnostic, and service utilization data on all patients seen by the service (Tansella et al., 1998).

Study design

The study was conducted with a longitudinal prospective design in the context of the South Verona Outcome Project (SVOP), a naturalistic longitudinal study that aimed to assess the outcome of care provided by the South Verona CMHS by using a set of standardized instruments completed in the frame of the routine clinical practice (Ruggeri et al., 1998a,b, 2001, 2002; Lasalvia et al., 2000, 2002).

In the SVOP, standardized assessments of patients took place twice a year, from April to June and from October to December. During these periods both new patients and patients already in contact with the service are assessed at the first or, at latest, the second time they saw a psychiatrist or a psychologist (except when the contact took place in the Casualty Department or in the Liaison Psychiatry Department).

For the present study, the three-month treated prevalence cohort attending the CMHS between 1 October and 31 December 1994 was assessed at two time points: at baseline and at 2 years.

Measures

Both baseline at follow-up assessments were performed by using a set of well-known standardized instruments. All patients were only interviewed after informed consent had been gained in each case. The procedure used was that research staff explained the purpose of the study and gave full details to each patient in writing, making clear that participation was entirely voluntary. We told potential subjects that they could choose whether to agree to participate, to decline or to agree and then withdraw at a

later time, without any detriment to their clinical care. Confidentiality was fully preserved.

Global Assessment of Functioning Scale (GAF) (Endicott et al., 1976; APA, 1994)

This is a measure of individual functioning on a continuous scale, where 0 denotes extremely severe dysfunction and 100 extremely good function.

Brief Psychiatric Rating Scale, 'expanded version' (BPRS) (Lukoff et al., 1986; Ventura et al., 1993)

The BPRS is among the most widely used instruments for evaluating symptom levels in psychiatric research; it is a semi-structured interviewer-administered scale that consists of 24 items rated on a seven-point severity scale, ranging from 1 ('symptom not present') to 7 ('extremely severe symptom'). Analyses were conducted using BPRS total mean score.

Disability Assessment Schedule-II (DAS) (WHO, 1988)

The DAS is a semi-structured interviewer-administered scale that assesses social disability in psychiatric patients. For this study, eight items from the section on social roles (participation in household activities, marital role, parental role, attrition in social contacts, work performance, interest in work, interest/information, and behaviour during emergencies) were used. Each item was rated on a six-point Likert scale (0 = no dysfunction; 5 = maximum dysfunction). Analyses were conducted using DAS total mean score.

Lancashire Quality of Life Profile (LQL) (Oliver, 1991)

The LQL is a self-rated questionnaire that assesses patients' satisfaction with life in 10 major life domains (global wellbeing, work/education, leisure, religion, finances, living situation, legal/safety, family relations, social relations and health). Subjective satisfaction ratings in each domain are reported on a seven-point Likert scale, ranging from 1 ('my life could not be worse') to 7 ('my life could not be better'). The LQL also includes objective measures in each domain, but these were not used in this study. Analyses were conducted on mean LQL total score.

Verona Service Satisfaction Scale (VSSS) (Ruggeri and Dall'Agnola, 1993)

The VSSS is a structured self-rated questionnaire assessing patients' satisfaction with the care received, which consists of 54 items covering seven dimensions of the patient's experience of mental health services in the previous year (overall satisfaction, professionals' skills and behaviour, information, access, efficacy, type of intervention, and relative's involvement). Items are rated on a five-point Likert scale (1 = terrible; 5 = excellent). Analyses were conducted on mean VSSS total score.

All instruments were the official Italian versions whose psychometric properties have been shown to be acceptable (BPRS: Roncone et al., 1999; DAS: Ruggeri and Nicolau, 2000; LQL: Lasalvia et al., 2000). Clinicians involved in the assessments were trained in the correct use of these standardized instruments. Inter-rater reliability for BPRS scores was assessed during the project and was always higher than 0.70 (Intra-Class Correlation Coefficient). In the case of self-rated scales, test-retest was performed in a sub-sample of subjects and showed good levels of stability (Ruggeri et al., 2001).

Total direct costs

Costs were attached to each service contact recorded in the PCR so as to give the best local estimates of long-run marginal opportunity costs (Beecham, 1995). Included costs relate to contacts with specialist mental health services (both public and private), while costs of care provided by general practitioners, private psychiatrists and psychologists, medication payments by patients, and all indirect costs were not considered in this study. However, from other studies we know that, among attenders of the South Verona CPS, less than 1% of patients receive private psychiatric ambulatory care (Amaddeo et al., unpublished manuscript). Individual total direct daily costs for the period between wave-B 1994 and 1996 follow-up assessments were calculated for each patient using PCR data, the unit cost list and a bespoke software package, linked to the Case Register and designed in Verona by our research group (Amaddeo et al., 1997a,b, 1998). In this study no attempt was made to separate inpatient and outpatient costs in the final model because only a minority of patients (43 out of

178) had been hospitalized during the index period. Costs are expressed in Euro.

Demographic, service utilization and diagnostic data were extracted from the South Verona PCR. Diagnoses were made according to ICD-10 codes and grouped in six categories according to the classification system used in the South Verona PCR: schizophrenia, severe affective disorders, depression without psychotic symptoms, neurotic and somatoform disorders, personality disorders and other diagnoses.¹ The clinicians making the diagnoses were specially trained in ICD-10 coding. An exercise performed with the staff involving four European case register centres demonstrated this system of categorization to be reliable across centres (Sytema et al., 1989). In some analyses, a dichotomous classification of 'psychotic' (including the PCR diagnosis of schizophrenia and severe affective disorders) versus 'non psychotic' (including all other PCR diagnoses) was used.

Service utilization data included the number, type and place of outpatient contacts, inpatient admissions and day-care attendance.

Statistical methods

To analyse the multivariate longitudinal dependencies among variables we used graphical Gaussian chain models for continuous data, which depict the structure of conditional associations by means of the partial correlation matrix. Graphical modelling is defined by imposing a simpler structure of association in which non-significant partial correlations are set to zero. A graphical Gaussian model has a one-to-one correspondence with a graph, where each vertex denotes a variable; the presence of an edge between a pair of vertices indicates a conditional association between the corresponding variables, so the lack of an edge indicates conditional independence. In general, graphical chain modelling is an extension of the graphical method, which is convenient when the variables may be grouped in a logical or temporal sequence set. More details on the graphical chain method are reported elsewhere (see Edwards, 2000

for a textbook; Ruggeri et al., 1998a for an application to psychiatric data).

Our analyses involved four stages. First, since our variables could be grouped in logically homogeneous blocks of explanatory, intermediate and response variables, we identified four sets of variables, based on a conceptualization tested in previous papers (Ruggeri et al., 1998a). *Set one* comprised disease-related characteristics such as diagnosis (dichotomized as psychotic/not psychotic) and illness duration at entry. *Set two* comprised the baseline evaluations of GAF, BPRS, DAS, LQL and VSSS (as predictors). *Set three* includes the total costs of patient care over two years (which were considered as an intermediate process variable). *Set four* comprised GAF, BPRS, DAS, LQL, and VSSS follow-up evaluations as outcome variables. Second, the associations between variables in set one were fixed by design, and the relationships among set one and set two variables were modelled. Third, a regression on costs was made using set 1 and set 2 variables as potential predictors. Finally, conditional dependencies among the set 4 response variables were analysed, with all the variables in the preceding sets fixed by design. A regression model for the outcome variables was performed, in order to make inferences regarding which variables of the preceding sets were predictors.

Model assumptions were checked using the methods outlined by Cox and Wermuth (1994), and absence of non-linear relationships and interactions was specifically tested. Logarithmic transformation was applied to daily costs, and square root transformation to illness duration. Since all variables considered are assumed being on a continuous scale, sensitivity analysis to departure from normality was conducted using the Box-Cox transformations. Model fitting, model selection and parameters estimation were performed using MIMWIN 2.3 (Edwards, 2000). Maximum likelihood estimates and likelihood ratios to compare nested models (backward elimination) were used (for details see Whittaker, 1990; Edwards, 2000; Cox and Wermuth, 1996); standard errors of the estimated regression coefficients were computed following Roverato and Whittaker (1993). Similar

¹Schizophrenia includes the following ICD-10 diagnoses: F20, F21, F22, F23, F24, F25, F28, F29, F84; severe affective disorders includes: F30, F31, F32.2, F33.3; depression without psychotic symptoms includes: F32 (.0, .1, .2, .8, .9), F33 (.0, .1, .2, .8, .9), F34.1, F41..2, F43 (.20, .21, 22); neurotic and somatoform disorders includes: F40, F41 (.0, .1, .3, .8, .9), F42, F44, F45, F48, F54; personality disorders includes: F34, F52, F60, F61, F62, F63, F64, F65, F66, F68, F69.

results were obtained using the above mentioned strategies, applied separately by two independent statisticians. For the regression analyses, Stata Release 6 was used (StataCorp, 1999).

Results

Characteristics of the sample

A sample of 261 patients completed the entire set of the study instruments (BPRS, GAF, DAS, LQL and VSSS) at baseline.

After 2 years, 12 patients from the baseline cohort had died, 15 were excluded due to the development of severe psychopathology and four could not be located; 230 were contacted and, of these, 178 (78%) completed the assessment at the 2-year follow-up. Thirty-six per cent of the patients were not in contact with the South Verona CMHS in the 6 months before the 2-year follow-up assessment and were considered to be no longer in contact with the service.

Responders were older than non-responders (46.9 years old versus 42.6, $p < 0.05$, Mann-Whitney U Test), more likely to be single (43.2% versus 12.7%, $p < 0.05$, chi-square test) and had a longer average time period since first contact with service (6.5 versus 4.8, $p < 0.01$, Mann-Whitney U Test).

Of patients assessed at follow-up, 62% were females, their mean age was 46.8 (range 18-86), 31% were employed, 39% were married, and 76% had up to 8 years of schooling. Regarding the diagnostic characteristics, 28% of the sample had a PCR diagnosis of schizophrenia, 11%, of severe affective disorders, 33% of depression without psychotic symptoms, 12% of neurotic and somatoform disorders, 11% of personality disorders and 5% other diagnoses. With regard to service utilization, 90% of patients had no previous hospitalization in 1994, 16% had been in contact with the service for less than one year, 23% for 1-5 years, 12% for 6-10 years, 49% for more than 10 years. The mean number of outpatient contacts in the year preceding the baseline assessment was 45.4 (SD 91.3), the days of voluntary admission 5.82 (SD 24.4), and the days of admission in sheltered apartments 9.25 (SD 55.1).

The mean interval between assessments was 28.3 months (SD 5.0, range 21.6-43.4).

The mean daily costs were 13,85 Euro (SD 33,54, median 1,64, range 0,05-189,61). Average total cost per patient in the follow-up period was 10778,53

Euro (SD 25398,28, median 1343,98, range 42,98-14751,85). Of these, 40.1% were inpatient costs (mean 4318,09, SD 16112,69, median 0, range 0-13138,56), and the remainder outpatient costs (mean 6460,44, SD 15059,73, median 1272,40, range 42,86-12799,66). Summary statistics for the variables considered in the analyses are given in Table 1.

Over the follow-up period, 25% of subjects had improved in the global functioning (GAF), whereas 44% had worsened their condition; 19% had higher level of disability (DAS) at follow-up with respect to baseline (versus 26% with a lower disability); 13% showed an improvement in levels of psychopathology (BPRS), while 12% worsened. Improvements in satisfaction with services (VSSS) and subjective quality of life (LQL) were found, respectively, in 15% and 26% of patients, while worsening in 21% and 19% respectively. Minimum detectable change in each instrument score has been assumed to be equal to $1/k$ (where k is the maximum possible score for each instrument) with the exception of GAF whose minimum detectable change, to maintain congruence with the Likert points of the other scales has been assumed to be $10/k$.

Relationships between set 1 and 2 variables

Table 2 shows the pattern of relationships between set 1 and 2 variables, and demonstrates that marginal and partial correlations give a different kind of information: some non-negligible marginal correlations vanish when analysed with partial correlations, such as the correlation between diagnosis and BPRS, and the correlation between GAF and VSSS. After having fitted a model that excluded non-significant partial correlations, diagnosis of psychosis was related only to illness duration and DAS. By contrast, GAF, BPRS and DAS were related each other. These variables were related to VSSS through DAS, whereas LQL was only related to VSSS.

Variables associated with direct costs

Only BPRS, diagnosis and illness duration were significant and positively associated with the logarithm of daily costs. As shown in Table 3, the more severe the psychopathology at baseline, the higher the patient costs in the follow-up period. Higher costs were also found to be predicted by a longer illness duration and a diagnosis of psychosis.

Table 1. Summary statistics of the twelve variables analysed on 178 cases assessed at baseline and 2-year follow-up

Variable	Mean or %	SD	Range
Set one:			
Psychosis	38.8%		
Illness duration	12.61	11.10	0–47.48
Square root(duration)	3.14	3.12	0–6.89
Set two:			
GAF baseline	61.60	15.70	20.00–90.00
BPRS baseline	1.46	0.41	1.00–3.71
DAS baseline	0.71	0.80	0.00–4.00
LQL baseline	4.46	0.88	1.92–6.12
VSSS baseline	3.96	0.54	2.15–5.00
Set three:			
daily cost	13.85	33.54	0.05–185.61
Set four:			
GAF follow-up	58.39	14.76	20.00–88.00
BPRS follow-up	1.48	0.42	1.00–3.42
DAS follow-up	0.64	0.87	0.00–4.33
LQL follow-up	4.58	0.88	1.68–6.17
VSSS follow-up	3.90	0.49	2.33–4.80

Variables associated with outcome

After the selection process, the likelihood ratio for the final model was 44.25 (38 d.f.), denoting a good fit. The regression coefficients of the predictors of each response variable are reported in Table 4. The GAF at follow-up was positively associated with its baseline level and with diagnosis (a diagnosis of psychosis implying lower global functioning at follow-up). Lower psychopathology at follow-up was predicted by lower global functioning at the baseline; this indicates that there is a greater chance of improvement for those more severely affected and that psychopathology is likely to be the patient characteristic more sensitive to care provided by health services in a 2-year span. The DAS and VSSS at follow-up are directly predicted by their respective baseline levels. Higher satisfaction with quality of life at follow-up was predicted by its baseline level and by baseline GAF. Patients with lower functioning at baseline were more satisfied with their life at follow-up.

Relationships between outcome variables

Marginal correlations, partial correlations and esti-

mated partial correlations among outcome variables are shown in Table 5. The GAF, DAS and BPRS were related each other, and indirectly correlated with VSSS through GAF.

Graphical modelling

The graphical Gaussian chain model produced is shown in Figure 1. In panel I the associations among set 1 and 2 variables (disease-related characteristics of patients and wave B 1994) are shown, and in panel III between set 3 variables (wave B 1996). Panel II shows the dependencies between cost and variables from sets 1 and 2, and panel IV shows the dependencies between all the variables in the sets 1 and 3 and the outcome variables (set 4). From the figure it is clear that the results described in sections 2 to 5 are easier to interpret when represented graphically.

Sensitivity analysis of the change between baseline and follow-up

Linearity in the relationships between pairs of variables and second-order interactions were checked following the methods described in the method section above (results not shown). More attention

Table 2. Marginal correlations, partial correlations, estimated partial correlations, p value and LR (right top triangle) between set 1 (disease-related characteristics) and set 2 (baseline variables)

Marginal correlations							
	Diagnosis	Illness duration	GAF	DAS	BPRS	LQL	VSSS
Diagnosis	1						
Illness duration	0.229	1					
GAF	-0.269	-0.116	1				
DAS	0.318	0.202	-0.743	1			
BPRS	0.222	0.104	-0.728	0.713	1		
LQL	0.055	0.037	0.087	-0.098	-0.093	1	
VSSS	-0.055	-0.042	0.168	-0.213	-0.095	0.512	1
Partial correlations							
	Diagnosis	Illness duration	GAF	DAS	BPRS	LQL	VSSS
Diagnosis	1						
Illness duration	0.175	1					
GAF	-0.065	0.042	1				
DAS	0.149	0.141	-0.430	1			
BPRS	-0.020	-0.031	-0.425	0.386	1		
LQL	0.086	0.047	-0.021	0.029	-0.086	1	
VSSS	-0.029	-0.023	0.059	-0.152	0.129	0.512	1
Estimated partial correlations							
	Diagnosis	Illness duration	GAF	DAS	BPRS	LQL	VSSS
Diagnosis	1						
Illness duration	0.218	1					
GAF	0	0	1				
DAS	0.197	0	-0.452	1			
BPRS	0	0	-0.423	0.364	1		
LQL	0	0	0	0	0	1	
VSSS	0	0	0	-0.113	0	0.504	1
Likelihood ratio test (bottom left: χ^2 with 1 d.f., upper right: p values)							
	Diagnosis	Illness duration	GAF	DAS	BPRS	LQL	VSSS
Diagnosis	–	0.002	–	<0.001	–	–	–
Illness duration	9.6	–	–	–	–	–	–
GAF	–	–	–	<0.001	<0.001	–	–
DAS	18.996	–	43.46	–	<0.001	–	0.004
BPRS	–	–	35.02	27.06	–	–	–
LQL	–	–	–	–	–	–	<0.001
VSSS	–	–	–	8.26	–	54.21	–

was dedicated to fit models to transformed data and checking the stability of the results. All variables (with the exception of diagnosis) were transformed using Box-Cox transformation and the analyses

have been repeated on the new data. Overall the results were highly consistent. A third analysis was conducted on derived variables: the degree of change (the difference between baseline and follow

Table 3. Regression coefficients of the relationships between logarithm of daily costs (set 3) with set 1 (disease-related characteristics) and set 2 (baseline variables); only significant predictors were included in the model (goodness-of-fit LR = 4.761, d.f. = 4, p = 0.3127)

Variable	b	SE(b)	t	Sig t
BPRS	1.612	0.277	5.819	<0.001
Diagnosis	1.423	0.235	6.046	<0.001
Illness duration (sqrt)	0.152	0.068	2.253	0.025
(Constant)	-1.856	0.442	(-4.195)	<0.001

Table 4. Adjusted regression coefficients of the predictors of response variables GAF FOLLOW-UP, BPRS FOLLOW-UP, DAS FOLLOW-UP, VSSS FOLLOW-UP and LQL FOLLOW-UP (goodness-of-fit LR = 44.2516, d.f. = 38, p = 0.2654)**4.1 Response: GAF follow-up**

Variable	Coeff.	Std. Err.	t	p
GAF baseline	0.231	0.0477	4.83979	<0.01
Diagnosis (0 non psychotic, 1 psychotic)	-3.2308	1.0322	-3.13008	0.0020

4.2 Response: BPRS follow-up

Variable	Coeff.	Std. Err.	t	p
GAF baseline	0.0031	0.00133	2.3334	0.0207

4.3 Response: DAS follow-up

Variable	Coeff.	Std. Err.	t	p
DAS baseline	0.2298	0.0522	4.4043	<0.001

4.4 Response: VSSS follow-up

Variable	Coeff.	Std. Err.	t	p
VSSS baseline	0.401	0.05816	6.8883	<0.001

4.5 Response: LQL follow-up

Variable	Coeff.	Std. Err.	t	p
LQL baseline	0.5467	0.0626	8.7304	<0.001
GAF baseline	-0.007	0.0029	-2.2429	0.0261

up) and the patient's average score at baseline and reassessment of the GAF, BPRS, DAS, VSSS, LQL. The aim was to identify whether the absolute level of a patient's score on a given variable affected the dependencies with other variables. In this analysis the graphical chain model includes duration of illness, diagnosis and average scores as predictors in

the first box, costs as intermediate variables in the second box, individual changes as response variables in the third box. The best fitting model confirmed the results previously obtained, showing the same pattern of conditional relationships among outcome indicators. However the lower psychopathology and greater satisfaction with life

Table 5. Marginal correlations, partial correlations and estimated partial correlations (with likelihood ratio tests – 1 degree of freedom – and p values) among set 4 (response variables) (model goodness-of-fit LR = 44.2516, d.f. = 38, p = 0.2654)

Marginal correlations					
	GAF Follow-up	DAS Follow-up	BPRS Follow-up	LQL Follow-up	VSSS Follow-up
GAF Follow-up	1				
DAS Follow-up	-0.535	1			
BPRS Follow-up	-0.659	0.642	1		
LQL Follow-up	0.147	-0.090	-0.099	1	
VSSS Follow-up	-0.074	-0.204	-0.028	0.401	1
Partial correlations					
	GAF Follow-up	DAS Follow-up	BPRS Follow-up	LQL Follow-up	VSSS Follow-up
GAF Follow-up	1				
DAS Follow-up	-0.153	1			
BPRS Follow-up	-0.495	0.487	1		
LQL Follow-up	0.055	0.090	-0.081	1	
VSSS Follow-up	0.098	-0.242	0.204	0.402	1
Estimated partial correlations					
	GAF Follow-up	DAS Follow-up	BPRS Follow-up	LQL Follow-up	VSSS Follow-up
GAF Follow-up	1				
DAS Follow-up	-0.187	1			
BPRS Follow-up	-0.504	0.437	1		
LQL Follow-up	0	0	0	1	
VSSS Follow-up	0.124	0	0	0.282	1
Likelihood ratio test (bottom left: χ^2 with 1 df, upper right: p values)					
	GAF Follow-up	DAS Follow-up	BPRS Follow-up	LQL Follow-up	VSSS Follow-up
GAF Follow-up	–	<0.0056	<0.001	–	0.001
DAS Follow-up	7.67	–	<0.001	–	–
BPRS Follow-up	56.61	40.46	–	–	–
LQL Follow-up	–	–	–	–	<0.001
VSSS Follow-up	10.85	–	–	54.21	–

at follow up predicted by lower global functioning at the baseline were no more statistically significant. Overall no predictors of differential changes reached statistical significance from this analysis.

Discussion

This is among the very first papers in mental health

research to apply the method of graphical chain models to routinely collected outcome data. It represents the extension of our previous 6-month longitudinal study on the outcome of psychiatric care (Ruggeri et al., 1998a): data cover a longer follow-up period (2 years versus 6 months); a larger set of outcome variables was used and disease-related

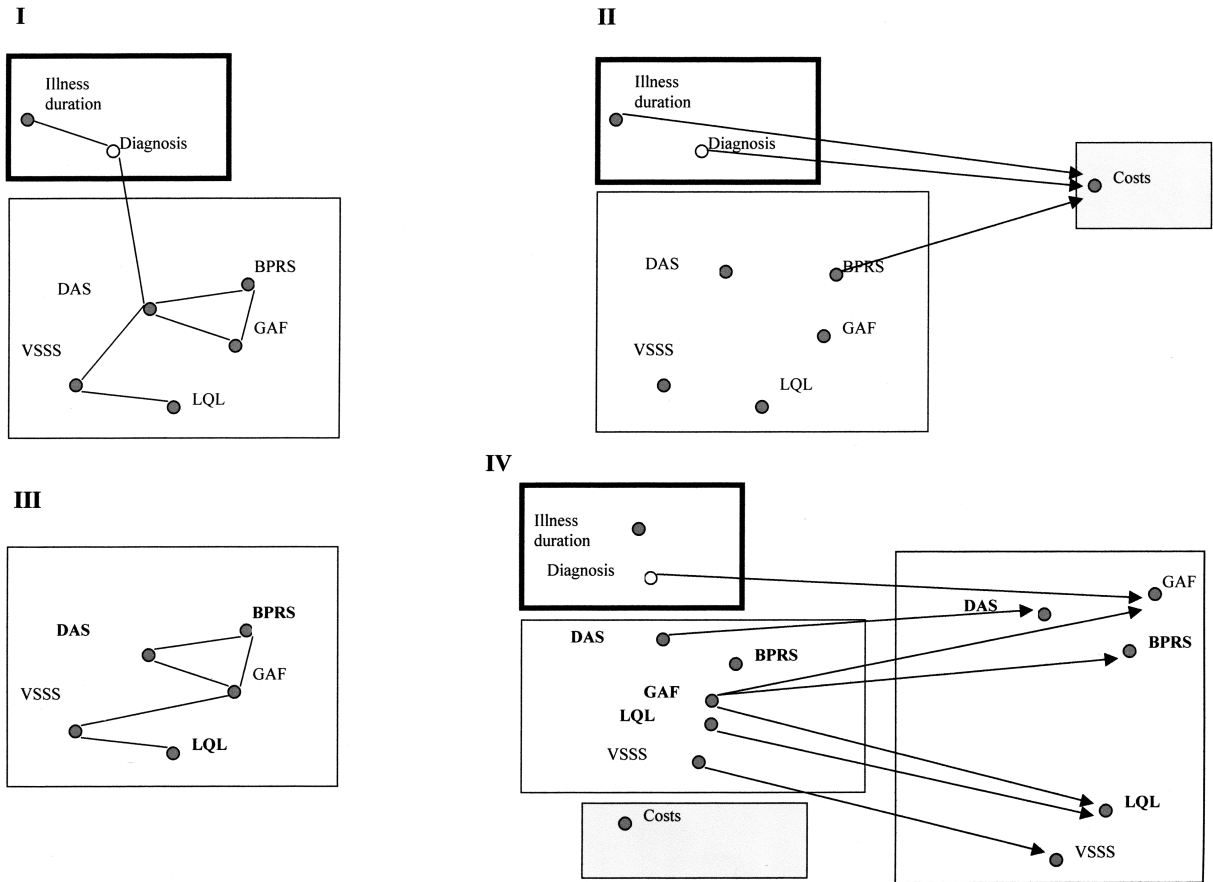


Figure 1. Graphs of the multivariate dependencies among variables.

characteristic of the patients, such as diagnosis and illness duration, have also been taken into account as predictors.

The present paper assessed:

- the relationships among outcome indicators;
- the effect of both disease-related characteristics and baseline measures of a series of outcome indicators on total costs of care in the follow up; and
- the effect of baseline indicators and costs on follow-up values of the outcome indicators.

The outcome variables assessed in this study showed the tendency to segregate, both at baseline and follow-up, in two distinct groups: a *clinical dimension*, which includes staff-rated measures such as global functioning, psychopathology and disability, and a *subjective dimension*, which includes patient-rated

measures such as satisfaction with services and quality of life. The tendency for a polarization of outcome measures in two distinct groups, already detected at 6 months (Ruggeri et al., 1998a), appeared to be *replicable over time*, as shown by the two-year follow-up data reported in the present study.

Cross-sectional associations among outcome indicators

Global functioning, psychopathology and social disability have been shown to be highly related to each other, thus confirming findings obtained in other studies (Anthony et al., 1995; Patterson et al., 1997; Hoffmann and Kupper, 1997; Ruggeri et al., 1998a; Judd et al., 2000).

Quality of life and service satisfaction are also strongly correlated with each other. In our data the relationships of subjective quality of life with the other

variables are mediated by satisfaction with services: a direct correlation between the two variables was found both at baseline and at the two-year follow-up, replicating previous findings (Ruggeri et al., 1998a, 2001, 2002). Two hypotheses seem plausible:

- that both satisfaction with services and quality of life reflect a common attitude of the patient;
- that in psychiatric patients quality of life is strongly determined by satisfaction with the care they received.

The first hypothesis has been recently tested by exploring the impact of personality traits over service satisfaction and satisfaction with life in a sample of patients with the full range of psychiatric conditions and followed by the South-Verona community base mental health service (Ruggeri et al., 2003a); the study showed that, although patients with traits of neuroticism were more dissatisfied with their lives, they were not more dissatisfied with the care they received. The second hypothesis has been explored in a recent European collaborative multi-site study (Ruggeri et al., 2003b), which showed cross-cultural stability of the association between service satisfaction and subjective quality of life and its predominance over a wide range of other explanatory variables, offering the appealing perspective that improvements in patients' subjective quality of life can be achieved by providing adequate and individualized care. However, for the time being, caution should be taken in interpreting these findings, as conflicting results have been obtained on possible confounders. On the one hand, previous papers have reported that subjective quality of life, self-rated needs and patients' assessment of treatments are all substantially correlated, thus suggesting the presence of an underlying subjective appraisal factor (Priebe et al., 1998; Fakhoury et al., 2002). On the other hand, it has been found that mood interferes with ratings of quality of life, but has a small influence on satisfaction with care (Atkinson and Caldwell, 1997). Achieving a better knowledge of these background aspects, and implementing follow-up studies on changes occurring over time in both quality of life and satisfaction with care might greatly contribute to a better understanding of this issue.

Service satisfaction is directly correlated to both social disability (baseline assessment) and global functioning (follow-up assessment): the less severe

the patients are, the more they are satisfied with services. Thus satisfaction with services seems to be more related to an underlying factor that can be interpreted as the capability of the patient to function appropriately in the society, but only indirectly related to psychopathology. As a matter of fact earlier studies have consistently found a weak relationship between service satisfaction and illness severity, as measured in terms of symptom levels (Ruggeri et al., 1998a, 2003b).

Costs and outcome indicators

Higher costs were predicted by higher psychopathology at baseline, diagnosis of a psychotic disorder and longer duration of illness. This finding is consistent with other studies (Amaddeo et al., 1997; McCrone et al., 1998; Bonizzato et al., 2000). In contrast to previous findings from the 6-month follow-up (Ruggeri et al., 1998a), no direct relationship was found between intervening costs and functioning at 2 years. To test whether this was because of the inclusion of variables not previously considered (diagnosis and duration of illness), data were re-analysed using only the previously included variables. In this analysis, cost was shown to be a predictor of severity at 2 years; this implies that cost of care reflects the overall pattern of care provided to the patients, which is mainly determined by diagnosis and duration of illness (Biggeri et al., 2001).

Predictors of outcome indicators

Our results show that baseline values for each variable are able to predict corresponding values at the two-years follow up. But lower psychopathology at follow-up was predicted only by lower global functioning at the baseline. Previous literature had reported similar results in various diagnostic groups (King et al., 1997; Links et al., 1998; Robinson et al., 1999); our study, however, investigated longitudinal associations for the first time using a multivariate approach, which enabled the indirect effects of the other variables in the model to be controlled. Table 1 shows that on average there is a tendency to stability for all outcome indicators. The graphical Gaussian chain model confirms that the most important predictor for each variable is the corresponding baseline level. Our results show that there is a greater chance of improvement in psychopathology for the more severe patients and that psychopathology is

likely to be the patient characteristic more sensitive to care provided by health services in a 2-year span. Higher functioning was indirectly correlated with higher satisfaction with life, cross-sectionally, and low functioning at baseline was a predictor for high satisfaction with life at follow-up. This indicates that an improvement in satisfaction with life at follow-up is experienced by patients who at baseline were at a lower functioning level, and consequently are those who are more able to experience higher absolute levels of improvement.

At the same time, we have found that a sizeable proportion of patients in the study change their status from baseline levels (ranging from over 60% for GAF to about 25% for BPRS) but these variations were not explained by the variables included in the model. Individual variations between reassessment and baseline levels might be explained by variables not considered in our study (such as socio-economical status, family and social environment, personality related variables) or by heterogeneity in the course of disorders.

Therefore, in a 2-year period there is little room for long-term trends in improvement or deterioration to become evident. Finally, the availability of only two assessments allows us to detect only linear changes between baseline and final assessment, while subjects who tend to vary could present strong non linear patterns which are missed by the two measurements taken at long time interval (Willet et al., 1998).

The finding that individual variations between reassessment and baseline levels are not predictable highlights one of the shortcomings of group-based ('nomothetic') enquiry. Individual differences will clearly account for some of the variation in a way that cannot be captured by investigating groups of patients. To fully understand the factors that might influence the outcome of care, high-quality group-based studies need to be integrated with individual ('idiographic') methods of enquiry. For example, this might involve use of qualitative techniques including outliers from the sample, to identify why some patients deteriorated whereas others improved from similar baseline levels.

The patients enrolled in this study may be considered representative of patients attending the South Verona CMHS in the study periods, who, in turn, have been shown to be representative of all south Verona residents receiving psychiatric care

(Balestrieri et al., 1994). The exclusion of patients who had very few contacts with the CMHS in the period should not have seriously biased the results, because these patients are low, and even occasional, users of services. The bias deriving from the higher severity of dysfunction of patients refusing to complete the quality of life and the satisfaction with services assessment is more relevant, but may be unavoidable in a naturalistic study of this type.

In conclusion, this study throws some light on the complex relationships between clinical, social and economic variables affecting the outcome of mental health care in the medium term. Our data show that, in the medium term, improvements in psychopathology, disability and global functioning are to be expected in the more severely ill; quality of life and service satisfaction are very strongly correlated and depend on baseline disability and improvement in functioning but not on psychopathology or total costs; the more ill get most support. While a substantial stability emerged on average, individual trajectories appeared to be sensitive to a complex set of variables. These findings are useful in order to formulate hypotheses on the outcome of mental health care, which should be tested in future studies.

Acknowledgements

The authors are grateful to Dr Doriana Cristofalo for assistance in data management. We are also grateful to the colleagues and the patients who participated in the South Verona Outcome Project, and specifically thank Rosa Dall'Agnola, Paola Bonizzato, Marco Stegagno, Giuseppe Imperadore, Nazario Santolini, Stylianos Nicolaou, Manuela Benetollo for their contribution in the data collection.

We are indebted to Professor Michele Tansella for his generous and continuous support and valuable advice. We thank Dr Mike Slade for having helped revise the manuscript.

This study was supported by the Istituto Superiore di Sanità (ISS), Roma, Progetto Nazionale Salute Mentale, with Grant (No. 96/QT/50) to Professor M. Ruggeri.

References

- Amaddeo F, Beecham J, Bonizzato P, Fenjo A, Knapp MJ, Tansella M. The use of a case register to evaluate the costs of psychiatric care. *Acta Psychiatr Scand* 1997a; 95: 189-98.
- Amaddeo F, Beecham J, Bonizzato P, Fenjo A, Tansella M, Knapp M. The costs of community-based psychiatric care for first-ever patients. A case register study.

- Psychol Med 1998; 28: 173–83.
- Amaddeo F, Bonizzato P, Tansella M. Psychiatric case registers for monitoring service utilization and evaluating its costs. In M Tansella ed Making Rational Mental Health Services. Roma: Il Pensiero Scientifico Editore, 1997b.
- American Psychiatric Association. DSM-IV. Diagnostic and Statistical Manual of Mental Disorders. 4 edn. Washington DC: APA, 1994.
- Anthony WA, Rogers ES, Cohen M, Davies RR. Relationships between psychiatric symptomatology, work skills, and future vocational performance. Psychiatr Serv 1995; 46: 353–8.
- Atkinson MJ, Caldwell L. The differential effects of mood on patients' ratings of life quality and satisfaction with their care. J Affect Disord 1997; 44: 169–75.
- Attkisson C, Cook J, Karno M, Lehman A, McGlashan H., Meltzer HJ, O'Connor M, Richardson T, Rosenblatt A, Wells K, Williams J, Hohmann AA. Clinical services research. Schizophr Bull 1992; 18: 561–626.
- Balestrieri M, Bon MG, Rodriguez-Sacristan A, Tansella M. Pathways to psychiatric care in south-Verona, Italy. Psychol Med 1994; 24: 641–9.
- Beecham J. Collecting and estimating costs. In Knapp M (ed.) The Economic Evaluation of Mental Health Care. Adelshot: Arena, 1995.
- Biggeri A, Bisoffi G, Rucci P, Ruggeri M, Tansella M. Graphical models for the multidimensional assessment of outcome. In Tansella M, Thornicroft G (eds) Mental Health Outcome Measures, Second Edition. London: Gaskell, Royal College of Psychiatrists, 2001.
- Bonizzato P, Bisoffi G, Amaddeo F, Chisholm D, Tansella M. Community-based mental health care: to what extent are service costs associated with clinical, social and service history variables? Psychol Med 2000; 30: 1205–15.
- Cox DR, Wermuth N. Multivariate Dependencies. Models, Analysis and Interpretation. London: Chapman & Hall, 1996.
- Cox DR, Wermuth N. Tests of linearity, multivariate normality and the adequacy of linear score. Applied Statistic 1994; 43: 347–55.
- Edwards D. Introduction to Graphical Modelling, 2 edn. New York: Springer-Verlag, 2000.
- Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976; 33: 766–71.
- Everitt BS. An Introduction to Latent Variable Models. London: Chapman & Hall, 1984.
- Fahrmeir L, Tutz G. Multivariate Statistical Modelling based on Generalized Linear Models. New York: Springer-Verlag, 1994.
- Fakhoury WK, Kaiser W, Roeder-Wanner UU, Priebe S. (2002). Subjective evaluation: is there more than one criterion? Schizophr Bull 2002; 28: 319–27.
- Hoffmann H, Kupper Z. Relationships between social competence, psychopathology and work performance and their predictive value for vocational rehabilitation of schizophrenic outpatients. Schizophr Res 1997; 23: 69–79.
- Jenkins R. Toward a system of outcome indicators for mental health care. Br J Psychiatry 1990; 157: 500–14.
- Jöreskog KG. Analysis of covariance structure. Scand J Stat 1981; 8: 65–92.
- Judd LL, Akiskal HS, Zeller PJ, Paulus M, Leon AC, Maser JD, Endicott J, Coryell W, Kunovac JL, Mueller TI, Rice JP, Keller MB. Psychosocial disability during the long-term course of unipolar major depressive disorder. Arch Gen Psychiatry 2000; 57: 375–80.
- Knudsen HC, Thornicroft M. Mental Health Service Evaluation. Cambridge: Cambridge University Press, 1996.
- Kreiner S. An informal introduction to graphical modelling. In Knudsen HC and Thornicroft G (eds) Mental Health Service Evaluation. Cambridge: Cambridge University Press, 1996.
- Lasalvia A, Ruggeri M, Mazzi MA, Dall'Agnola RB. The perception of needs for care in staff and patients in community-based mental health services. The South-Verona Outcome Project 3. Acta Psychiatr Scand 2000; 102: 366–75.
- Lasalvia A, Ruggeri M, Dall'Agnola RB. Lancashire Quality of Life Profile. In Ruggeri M, Dall'Agnola RB. Come Valutare l'Esito nei Dipartimenti di Salute Mentale. Roma: Il Pensiero Scientifico Editore, 2000.
- Lasalvia A, Ruggeri M, Santolini N. Subjective quality of life: its relationship with clinician-rated and patient-rated psychopathology. The South-Verona Outcome Project 6. Psychother Psychosom 2002; 71: 275–84.
- Lauritzen SL. Graphical Models. Oxford: Oxford University Press, 1996.
- Linder-Pelz S. Social psychological determinants of patient satisfaction: a test of five hypotheses. Soc Sci Med 1982; 16: 583–9.
- Links PS, Heslegrave R, Van Reekum R. Prospective follow-up study of borderline personality disorder: prognosis, prediction of outcome, and Axis II comorbidity. Can J Psychiatry 1998; 43: 265–70.
- Lukoff D, Liberman RP, Nuechterlein KH. Manual for the Expanded Brief Psychiatric Rating Scale. Schizophr Bull 1986; 12: 594–602.
- McCrone P, Thornicroft G, Parkman S, Nathaniel-James D, Ojorongbe W. Predictors of mental health service costs for representative cases of psychosis in south London. Psychol Med 1998; 28: 159–64.
- Norman RM, Malla AK, Cortese L, Cheng S, Diaz K, McIntosh E, McLean TS, Rickwood A, Voruganti LP. Symptoms and cognition as predictors of community functioning: a prospective analysis. Am J Psychiatry 1999; 156: 400–5.
- Oliver JP. The social care directive: development of a

- quality of life profile for use in community services for the mentally ill. *Soc Work Soc Sci Rev* 1991; 3: 5–45.
- Patterson TL, Semple SJ, Shaw WS, Halpain M, Moscona S, Grant I, Jeste DV. Self-reported social functioning among older patients with schizophrenia. *Schizophr Res* 1997; 27: 199–210.
- Priebe S, Kaiser W, Huxley PJ, Roder-Wanner UU, Rudolf H. Do different subjective evaluation criteria reflect distinct constructs? *J Nerv Ment Dis* 1998; 186: 385–92.
- Robinson DG, Woerner MG, Alvir JM, Geisler S, Koreen A, Sheitman B, Chakos M, Mayerhoff D, Bilder R, Goldman R, Lieberman JA. Predictors of treatment response from a first episode of schizophrenia or schizoaffective disorder. *Am J Psychiatry* 1999; 156: 544–9.
- Roncione R, Ventura J, Impallomeni M, Falloon IR, Morosini PL, Chiaravalle E, Casacchia M. Reliability of an Italian standardized and expanded Brief Psychiatric Rating Scale (BPRS 4.0) in raters with high vs. low clinical experience. *Acta Psychiatr Scand* 1999; 100: 229–36.
- Roverato A, Whittaker J. Standard Errors for the Parameters of Graphical Gaussian Models. Technical Report. Department of Mathematics, University of Lancaster (UK), 1993.
- Ruggeri M, Dall'Agnola R. The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients, relatives and professionals. *Psychol Med* 1993; 23: 511–23.
- Ruggeri M, Tansella M. Evaluating outcome in mental health care. *Curr Opin Psychiatry* 1995; 8: 116–21.
- Ruggeri M, Biggeri A, Rucci P, Tansella M. Multivariate analysis of outcome of mental health care using graphical chain models. The South-Verona Outcome Project 1. *Psychol Med*. 1998a; 28: 1421–31.
- Ruggeri M, Riani M, Rucci P, Biggeri A, Tansella M. Multidimensional assessment of outcome in psychiatry. The use of graphical displays. The South-Verona Outcome Project 2. *Int J Methods Psychiatr Res* 1998b; 7: 186–98.
- Ruggeri M, Bisoffi G, Fontecedro L, Warner R. Subjective and objective dimensions of quality of life in psychiatric patients: a factor analytical approach. The South-Verona Outcome Project 4. *Br J Psychiatry* 2001; 178: 268–75.
- Ruggeri M, Gater R, Bisoffi G, Barbui C, Tansella M. Determinants of subjective quality of life in patients attending community-based mental health services. The South-Verona Outcome Project 5. *Acta Psychiatr Scand*. 2002; 105: 131–40.
- Ruggeri M, Pacati P, Goldberg D. Neurotics are dissatisfied with life, but not with services. The South Verona Outcome Project 7. *Gen Hosp Psychiatry* 2003a; 25: 338–44.
- Ruggeri M, Lasalvia A, Bisoffi G, Thornicroft G, Vazquez-Barquero JL, Becker T, Knapp M, Knudsen HC, Schene A, Tansella M; EPSILON Study Group. Satisfaction with mental health services among people with schizophrenia in five European sites: results from the EPSILON Study. *Schizophr Bull* 2003b; 29: 229–45.
- StataCorp. Stata Statistical Software: Release 6.0. Stata Corporation: College Station TX, 1999.
- Sytema S, Giel R, Ten Horn GHMM, Balestrieri M. The reliability of diagnostic coding in psychiatric case registers. *Psychol Med* 1989; 19: 999–1006.
- Tansella M, Amaddeo F, Burti L, Ruggeri M. Community-based mental health care in Verona, Italy. In Goldberg D, Thornicroft G (eds) *Mental Health in our Future Cities* Hove: Psychology Press, 1998.
- Thornicroft G, Tansella M. The Mental Health Matrix. A Manual to Improve Services. Cambridge: Cambridge University Press, 1999.
- Ventura J, Green M, Shaner A, Liberman R. Training and quality assurance with the Brief Psychiatric Rating Scale: 'The drift busters'. *Int J Methods Psychiatr Res* 1993; 3: 221–44.
- Whittaker J. Graphical Models. Chichester: Wiley, 1990.
- World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: WHO, 1992.
- World Health Organization. Disability Assessment Schedule (DAS-II). Geneva: WHO, 1988.
- Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J Am Stat Assoc* 1962; 57: 348–68.

Correspondence: Professor Mirella Ruggeri, Dipartimento di Medicina e Sanità Pubblica, Sezione di Psichiatria, Ospedale Policlinico, P. le L.A. Scuro, 10 37134 Verona, Italia. Telephone (+39) 45-8074441. Fax (+39) 45-585871. Email: mirella.ruggeri@univr.it.