Multivariate analysis of outcome of mental health care using graphical chain models
The South-Verona Outcome Project 1

M. RUGGERI,[1] A. BIGGERI, P. RUCCI AND M. TANSELLA

From the Servizio di Psicologia Medica, Istituto di Psichiatria, Università di Verona, Verona; and Dipartimento di Statistica, Università di Firenze, Firenze, Italy

ABSTRACT

Background. Short-term outcome of mental health care was assessed in a multidimensional perspective using graphical chain models, a new multivariate method that analyses the relationship between variables conditionally, i.e. taking into account the effect of antecedent and intervening variables.

Methods. GAF, BPRS, DAS (at baseline and after 6 months), LQL and VSSS (at follow-up only) were administered to 194 patients attending the South-Verona community-based mental health service. Direct costs in the interval were also calculated. Graphical chain models were used to analyse: (1) the associations between predictors (psychopathology, disability, functioning, assessed at baseline); (2) the effects of predictors on costs; and (3) the effect of predictors and costs on outcomes (psychopathology, disability, functioning, quality of life and service satisfaction) as well as their correlation.

Results. Psychopathology, disability and functioning scores at baseline predicted the corresponding scores at 6-month follow-up, with greater improvement in the more severely ill. Higher psychopathology and poorer functioning at baseline predicted higher costs and, in turn, costs predicted poorer functioning at follow-up. Outcome indicators polarized in two groups: psychopathology, disability and functioning, which were highly correlated; and the dyad service satisfaction and quality of life. Service satisfaction was highly related to quality of life and was predicted by low disability and high dysfunctioning. No predictors for quality of life were found.

Conclusions. Graphical chain models were demonstrated to be a useful methodology to analyse process and outcome data. The results of the present study help in formulating specific hypotheses for future studies on outcome.

INTRODUCTION

Many authors have stated that the assessment of outcome in psychiatry should be multidimensional and that the effects of treatments should be evaluated using multiple measures that explore many areas, including psychopathology, social disability, quality of life, service satisfaction and service utilization (Schulberg & Bromet, 1981; Wright *et al.* 1989; Jenkins, 1990;

Mirin & Namerow, 1991; Attkisson *et al.* 1992; Ruggeri & Tansella, 1995, 1996; Knudsen & Thornicroft, 1996; Tansella & Ruggeri, 1996).

However, in mental health research, analyses are usually performed either by considering one outcome indicator at a time, or using multivariate analyses whose results, when many variables are taken into account, are difficult to interpret. In the last few years considerable advances have appeared in the statistical literature and graphical modelling has been advocated as a valuable tool for addressing research hypotheses in a multivariate data set (Lauritzen & Wermuth, 1989; Wermuth & Lauritzen, 1990;

[1] Address for correspondence: Dr Mirella Ruggeri, Servizio di Psicologia Medica, Istituto di Psichiatria, Ospedale Policlinico, 37134 Verona, Italy.

Whittaker, 1990; Edwards, 1995; Cox & Wermuth, 1996). The underlying idea is that 'any kind of relationship should be analysed as a *conditional* relationship' (Kreiner, 1996), that is, a relationship between two variables can be considered as proved only if it does not disappear after controlling for the effect of antecedent or intervening variables. In the present context this means that we cannot fully understand the determinants of outcome without taking into account the pattern of relationships among its multiple dimensions.[1]

Graphical chain models are a type of graphical model especially suited to analyse complex sets of variables, some of which are predictors, some responses, and some intermediate between them. The study of the undirected associations among such variables is first conducted within each set; directed causal association is then assessed conditionally to all associations among predictors and intermediate variables.

The aim of this paper is to investigate, by means of graphical chain models, the associations among three sets of variables; (*a*) three variables which can be considered as predictors of outcome and represent baseline individual measures of the corresponding outcome indicators – global functioning, psychopathology and social disability; (*b*) total costs of specialist care, a process variable, which summarizes all information on service utilization in a given interval; (*c*) a series of variables considered as the most useful indicators of the outcome of mental health care – global functioning, psychopathology, social disability, quality of life and satisfaction with services.

[1] This goal is not achievable by other methods: multivariate regression would give the same answers if the same set of predictors affect all responses but it does not allow modelling the conditional associations among responses and, even more important, the regression coefficients are not adjusted for the correlation among responses (Fahrmeir & Tutz, 1994); seemingly unrelated regressions (an econometric technique) take into account the correlation among responses but this is appropriate only if a different set of predictors are used for each response (Zellner, 1962); structural equation models (such as LISREL models, see Jöreskog, 1981) are not appropriate here, first, because circular relationships among the variables are allowed, and secondly, because it is not always possible to interpret the coefficients of the relationships represented in the LISREL models as regression coefficients, i.e. changes in the response(s) for unit change in the explanatory variable(s) keeping the other explanatory variables constant. Finally, other techniques rely on the assumption of a causal role of latent, non-observable, variables: while they can be encompassed in graphical modelling, we do not introduce them in the present application because of the substantial arbitrariness in their specification and the *post hoc* nature of their interpretation (Everitt, 1984).

Specifically, we aim to study, in a consecutive series of patients receiving care by a community-based mental health service (CMHS): (1) the associations between global functioning, psychopathology, social disability assessed at baseline (predictors); (2) the effects of predictors on total costs of care in the following 6 months; and (3) the effects of predictors and of total costs of care on all indicators of the 6-month outcome.

Data were collected within a naturalistic, longitudinal study aimed to assess the outcome of care provided by a CMHS (the South-Verona Outcome Project (SVOP)) (Ruggeri & Tansella, 1996; Tansella & Ruggeri, 1996). A preliminary analysis was reported in Biggeri *et al.* (1996). This paper is the first in a series that will describe the data collected in the SVOP.

METHOD

The area and the mental health services

Data were collected in South-Verona, a mainly urban area in north east of Italy, which is relatively affluent and predominantly middle class, with a low migration rate. The total population is about 75000 inhabitants. The South-Verona CMHS, established in 1978, 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 in-patient care, day-care, rehabilitation, out-patient care and home visits, as well as a 24 hour emergency service and residential facilities for long-term patients. The same staff (excluding nurses) works both in hospital and in community facilities (Tansella, 1991). It has been proved that this 'single staff module' ensures higher continuity of care over facilities and over time, as compared with a model of community care characterized by different staff working in and outside hospital (Sytema *et al.* 1997).

A Psychiatric Case Register (PCR), which covers the same geographical area (South-Verona), has been operating since 31 December 1978. Admissions to psychiatric wards, private clinics and neurological wards (with a psychiatric diagnosis) are recorded. Also the number of days spent in hostels run by the South-Verona CMHS are recorded. All mental health services providing care to South-Verona residents, including in-patient private institutions, out-

patient services for substance misuse and mental health services for children and adolescents, report to the PCR information regarding patients over 14 years of age. Psychiatrists and psychologists in private practice and general practitioners (GPs) do not report to the PCR. Out-patient contacts (all other contacts but admissions to hospital, hostels and wards) with psychiatrists, psychologists, social workers and psychiatric nurses are also recorded. Each attendance at an out-patient clinic and each domiciliary visit are counted as a contact. The same applies to day-care at day-hospital and centre, where each daily attendance is counted as one contact. For more details see Tansella (1991).

Patients

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

In wave A the assessment is made only by the key-professional on the basis of the patient's condition in the previous month and includes assessment of global functioning, psychopathology, and disability in performing social roles. In wave B the assessment is made both by the key-professionals (again the same assessments as in wave A) and by the patients, who are requested to assess their quality of life and satisfaction with mental health services in relation to the experience over the previous year.

Assessments

This study includes patients who have been assessed both in wave A and in wave B of 1994 assessments and uses the total mean scores of the following instruments.

Global Assessment of Functioning Scale (Endicott & Spitzer, 1976)

This is a measure of individual functioning on a continuous scale, where 0 denotes extremely severe dysfunction and 90 extremely good function.

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

This consists of 24 items rated on a 7-point Likert scale (1 = no symptom; 7 = extremely severe symptom) which conceptually cover five dimensions: anxiety/depression, positive symptoms, negative symptoms, mania and cognitive impairment.

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

Eight items from the section on Social Roles have been used. They cover the following roles: participation in household activities; marital role; parental role; attrition in social contacts; work performance; interest in work; interest/information; and behaviour during emergencies. Each item is rated on a 6-point Likert scale (0 = no dysfunction; 5 = maximum dysfunction).

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

This consists of 164 items covering 13 areas: general well-being (assessed in two sections); leisure/participation; religion; finances; living situation; legal and safety; family relations; social relations; health; and self-concept. Each section assesses separately objective and subjective quality of life: the former items are rated on a varying scale (these cannot be summarized in a global score and are not taken into account in this paper); the latter items are rated on a 7-point Likert scale (1 = my life couldn't be worse; 7 = my life couldn't be better).

Verona Service Satisfaction Scale (VSSS; Ruggeri and Dall'Agnola, 1993; Ruggeri *et al.* 1994, 1996)

This 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 involvement of the relative. Items are rated on a 5-point Likert scale (1 = terrible; 5 = excellent).

All versions of scales and questionnaires reported above are the official Italian versions. All key-professionals have been trained in the correct use of the standardized instruments.

Inter-rater reliability has been monitored and was always higher than 0·70. A research team helped the patients in filling in the instruments if needed and assessed their understanding of items and coherence of assessments; confidentiality was fully preserved.

Total direct costs

Costs were attached to each service contact recorded on the PCR so as to give the best local estimates of long-run marginal opportunity costs. This is the appropriate base for economic evaluations of this kind, and ensures that both capital and running costs are included, and that the long-term resource consequences of any changes to service utilization patterns are explicitly identified and measured (Beecham, 1995). The costs included are those concerning 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% gets private psychiatric ambulatory care (Amaddeo *et al.* unpublished manuscript).

Costs are expressed in Italian lire at 1993 price levels. The costs of support and treatment provided by the South-Verona CMHS were estimated by dividing the total expenditure in 1993 for each service component by the number of items of care in the same year, and included the costs of personnel (including time spent in clinical or planning meetings), administration, medication, examinations and accommodation (including capital costs). The cost of out-patient contacts were calculated taking into account the total cost of a working minute for the various professionals (psychiatrists, psychologists, social workers, nurses, etc.) and multiplying this value by the estimated time spent on each contact. For the in-patient services the cost per day was calculated and for the rehabilitation groups and the day-hospital a cost per contact was estimated. The costs of private in-patient clinics were based on the prices paid by the Italian National Health Service. The support and treatment services listed here are mainly delivered, co-ordinated, and funded by the public sector health service.

Individual total direct costs referring to the period between wave-A and wave-B 1994 assessments was 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.* 1997*a*, *b*, 1998). In this study no attempt was made to separate in-patient and out-patient costs in the final model since only a minority of patients (17 out of 194) were hospitalized during the index period.

Statistical methods

The partial correlation matrix depicted the structure of association among the nine continuous variables. Each element of the matrix denotes the association between two variables conditional to all the others. Zero partial correlation coefficients mean conditional independence between two variables. Graphical modelling aims to impose a simpler structure of association among variables, in which non-significant partial correlations are set to zero. The result of graphical modelling is a graph, where each vertex denotes a variable; the presence of an edge between a pair of vertices indicates association between the corresponding variables and the lack of an edge means, therefore, conditional independence. This does not mean that the two variables are marginally independent. They may be dependent if the two vertices are connected indirectly by other vertices.

Since our variables could be 'naturally' grouped in subsets of explanatory, intermediate and response variables, we built up a graphical chain model, in which the association between variables was constrained to follow their temporal sequence. Specifically, we proceeded as follows: first we identified three sets of variables: set one, baseline values of GAF, BPRS, and DAS, which are considered as predictors; set two, total cost of patient care, which is an intermediate process variable; set three, outcome values of BPRS, GAF, DAS, LQL, and VSSS, which together represent the outcome.

Then, we conducted the analysis in three steps: (1) we modelled the relationships among the baseline variables, without considering the other sets; (2) we identified which of set one variables affected the total direct costs; (3) we assessed which variables from sets one and two

predicted the multivariate response (i.e. functioning, psychopathological status and social disability at end point, quality of life and satisfaction). This last step implied simultaneous modelling of the associations among response variables.

The graphical chain model relies on the assumptions that non-linear relationships are negligible, and that a relationship between pairs of variables is not modified by a third variable, i.e. there are no significative interaction terms between pairs of predictors. We checked for any violation of these assumptions using the method of Cox & Wermuth (1994). Logarithmic transformations were applied to disability and total direct costs (A-DAS, B-DAS, total costs), which displayed a curvature in the bivariate relationships with other variables. Then, since all variables considered are assumed on a continuous scale, sensitivity analysis to departure from normality, with special consideration for asymmetry of the distribution was conducted using Box–Cox transformations. Since the results confirmed those obtained on the original scale we do not present this part of the analysis (available upon request). In conclusion, we can safely assume that all prerequisites of the model are met.

Estimation, fitting and selection of the best graphical model are performed using MIM (Edwards, 1993). Maximum likelihood estimates and likelihood ratios to compare nested models (backward elimination) are used (for details see Whittaker, 1990; Edwards, 1995); standard errors of the estimated regression coefficients are computed following Roverato & Whittaker (1993). For the regression analyses, SPSS for Windows 95, version 7.0, was used (SPSS Inc., 1996). For a thorough presentation of these methods see Cox & Wermuth (1996).

RESULTS

1 Patients and assessments

Three-hundred-fifty-two patients with at least one contact with a psychiatrist or a psychologist both in April–June and in October–December 1994 were initially selected. Out of these, 243 (69%) were assessed at both waves by a key-professional who completed GAF, BPRS and DAS. They were similar to non-assessed patients on demographic characteristics, diagnostic dis-

Table 1. *Summary statistics of the nine variables analysed on* 194 *cases assessed at a 6-month interval* (*wave A and B*)

Variable	Mean	(s.d.)	Range
A-GAF	59·57	16·01	10·00–90·00
A-BPRS	1·57	0·48	1·00–3·46
A-DAS	1·06	0·73	0·14–4·00
Log A-DAS	0·67	0·32	0·13–1·61
Total costs (lire)	4 108,113	10 726,037	130,790–100 568,170
Log-total costs	7·06	1·38	4·87–11·52
B-GAF	62·55	15·71	24·00–90·00
B-BPRS	1·46	0·41	1–3·71
B-DAS	0·68	0·78	0–3·83
Log B-DAS	0·43	0·41	0–1·58
LQL	4·48	0·86	1·88–6·16
VSSS	3·98	0·50	2·29–4·85

tribution and time to first psychiatric contact, but showed a significantly more intensive service utilization in the year before the index wave A assessment (out-patient contacts in the last year $40 \cdot 5 \pm 73 \cdot 0$ *v.* $25 \cdot 5 \pm 49 \cdot 5$, $P = 0 \cdot 002$, Mann–Whitney U test).

Out of the 243 patients assessed by professionals, 194 (80%) completed also the VSSS and LQL. These patients were similar in all respects to patients who did not complete the VSSS and LQL, except for a better functioning (mean GAF $59 \cdot 5 \pm 16 \cdot 1$ *v.* $53 \pm 17 \cdot 8$, $P = 0 \cdot 025$, Mann–Whitney U test).

Sociodemographic characteristics of the fully assessed patients are as follows: 63·4% were females, the mean age was 45·6 (range 19–82), 35·8% were employed, 43·5% were married, and 74·1% had up to 8 years of schooling. The diagnostic distribution, according to groupings of ICD-10 (WHO, 1992) diagnoses recorded in the PCR, was the following: schizophrenia and related disorders 27·5%; affective psychosis 10·9%; depressive neurosis 31·1%; other neurotic and somatoform disorders 13·5%; personality disorders 11·4%; other diagnoses 5·7%. As to service utilization, 86·5% had no previous hospitalization; 17·3% had been in contact with the service for less than 1 year; 37·2% for 1–5 years; 19·4% for 6–10 years; 26·2% for more than 10 years. The number of out-patient contacts in the year before wave A assessment was $40 \cdot 5 \pm 73 \cdot 6$, the days of voluntary admission $7 \cdot 95 \pm 27 \cdot 79$, the days of compulsory admissions $0 \cdot 11 \pm 1 \cdot 19$, and the days of admission in sheltered apartments $8 \cdot 02 \pm 51 \cdot 63$.

Table 2. *Correlations (left bottom triangle) and partial correlations (right top triangle) between set 1 variables**

	A-GAF	A-BPRS	A-DAS
A-GAF		−0·670	−0·188
A-BPRS	−0·774		0·296
A-DAS	−0·545	0·579	

* GAF-BPRS LR = 115·6, *P* < 0·001; GAF-DAS LR = 6·96, *P* = 0·008; BPRS-DAS LR = 17·8, *P* < 0·001.

Table 3. *Regression coefficients reflecting the relationships between total costs and baseline values of psychopathology, functioning and disability**

Variable	b	s.e.(b)	t	P
A-GAF	−0·017	0·008	−2·121	0·035
A-BPRS	0·979	0·281	3·484	< 0·001
log A-DAS	0·407	0·319	1·276	0·203
(Constant)	6·281	0·870	(7·216)	< 0·001

* Goodness-of-fit LR = 1·66, df = 1, *P* = 0·197.

The mean interval between wave A and wave B assessments was 5·19 months, s.d. 1·02, range 2·7–7·8. Average total costs per patient in this interval were equal to 4 108,113 Italian lire ± 10 726,036 (range 130,790–100 586,170). Of these 31·5 % were in-patient costs (mean 1 295,936 ± 5 975,196, range 0–45 283,560) and the remainder out-patient costs (mean 2 812,176 ± 6 595,814, range 130,790–65 819,650).

Summary statistics for the nine variables are given in Table 1.

2 Relationships between predictive variables

The pattern of relationships between set one variables was first investigated. Marginal correlations between the three variables (Table 2) were high, denoting a strong association between them. Partial correlations maintained their statistical significance but changed in magnitude: specifically, log-A-DAS showed the weakest correlation with the other variables.

3 Predictors of direct costs

The regression coefficients of GAF, BPRS and DAS on total costs are reported in Table 3. Costs are related to both psychopathology and functioning at baseline: as expected higher costs are predicted by poorer functioning and higher psychopathology, while also the relationship

Table 4. *Adjusted regression coefficients for the four response variables B-GAF, B-BPRS, B-DAS and VSSS and their predictors**

Variable	b	s.e.(b)	t	P
4.1 Response: B-GAF				
A-GAF	0·412	0·058	7·099	< 0·001
A-BPRS	9·412	1·876	5·016	< 0·001
Total costs	−1·176	0·343	−3·430	0·006
4.2 Response: B-BPRS				
A-GAF	0·007	0·002	3·943	< 0·001
A-BPRS	0·438	0·056	7·782	< 0·001
Log-A-DAS	−0·192	0·055	−3·516	< 0·001
4.3 Response: B-DAS				
Log A-DAS	0·258	0·055	4·680	< 0·001
4.4 Response: VSSS				
A-GAF	−0·006	0·003	−2·335	0·019
Log A-DAS	−0·306	0·100	−3·050	0·002

* Goodness-of-fit LR = 9·03, df = 16, *P* = 0·91.

with disability is direct, but not significant. The best fitting model had a marginal likelihood ratio value of 1·66 (1 df) and $r^2 = 0·33$; it retained A-GAF and A-BPRS as predictors of total costs.

4 Predictors of outcome

The best fitting model had a marginal likelihood ratio of 9·077 (16 df), denoting a good fit. Table 4 shows the adjusted regression coefficients from that model for each response separately.

We found that global functioning at wave B was predicted by its corresponding level on wave A, by BPRS score at wave A (higher severity at BPRS in wave A predicted lower dysfunctioning in wave B) and by total costs (higher costs higher dysfunctioning) (Table 4.1). Psychopathology at wave B was predicted by its corresponding level at wave A and by GAF score and DAS score at wave A (higher dysfunction and disability at wave A predicted lower psychopathology in wave B) (Table 4.2). Social disability at wave B was predicted only by its corresponding level at wave A (Table 4.3).

In the more severely ill, a tendency to an improvement of the baseline values of BPRS, GAF and DAS was found at follow-up. This result needs some explanation. In fact, if there were no improvements, the baseline and final levels would coincide, and therefore they could be displayed on a straight line with unit slope and zero intercept. If the subjects were to

Table 5. *Empirical correlations, partial correlations and estimated partial correlations (with* P *values) among the response variables*

Empirical correlations					
B-GAF	1				
B-BPRS	−0·760	1			
LQL	0·117	−0·100	1		
VSSS	0·184	−0·091	0·489	1	
B-DAS	−0·777	0·719	−0·150	−0·214	1
	B-GAF	B-BPRS	LQL	VSSS	B-DAS
Partial correlations					
B-GAF	1				
B-BPRS	−0·470	1			
LQL	−0·062	−0·047	1		
VSSS	0·128	0·115	0·471	1	
B-DAS	−0·400	0·291	−0·033	−0·070	1
	B-GAF	B-BPRS	LQL	VSSS	B-DAS
Estimated partial correlations					
B-GAF	1				
B-BPRS	−0·471	1			
LQL	0	0	1		
VSSS	0·102	0	0·472	1	
B-DAS	−0·405	0·273	0	0	1
	B-GAF	B-BPRS	LQL	VSSS	B-DAS
Likelihood ratio test (bottom left: χ^2 with 1 df, upper right: P values)					
B-GAF	1	< 0·0001	NS	0·0133	< 0·0001
B-BPRS	55·977	1	NS	NS	< 0·0001
LQL	0	0	1	< 0·0001	NS
VSSS	6·122	0	52·955	1	NS
B-DAS	48·958	20·378	0	0	1
	B-GAF	B-BPRS	LQL	VSSS	B-DAS

improve irrespective of their baseline level then, on average, the observations would be aligned on a straight line with unit slope but with the intercept different from zero, which can be interpreted as the mean improvement. Finally, if improvements are related to the baseline levels (the lower the baseline the greater the improvement), then the observations would be aligned on a straight line with a slope lower than unity. Our data show that the slope for psychopathology and global functioning was about 0·4 (Table 4.1 and 4.2 respectively), for social disability it was 0·25 (Table 4.3).

High satisfaction with services was predicted by high dysfunctioning, but also by low disability, at wave A (Table 4.4). No predictors for quality of life have been found.

5 Relationships between outcome indicators

Table 5 shows empirical, partial and fitted correlations among outcome indicators. Given the effect of predictors, we found that: (1) functioning, psychopathology, and disability at wave B are related to each other with the same pattern seen at wave A; (2) satisfaction with services is positively correlated with functioning at wave B and with quality of life, but is conditionally independent from all other wave B variables; (3) quality of life is positively correlated with satisfaction and conditionally independent from the rest of wave B variables.

6 Graphical modelling

The graphical chain model finally selected is shown in Fig. 1. The associations between wave A and wave B variables are displayed in panels I and III, respectively. The causal relationships between set one variables and costs are shown in panel II. The causal relationships between variables belonging to set one and set two and outcome variables are shown in panel IV. The steps taken to obtain this model are detailed in sections 2 to 5 of this Results section.

DISCUSSION

To our knowledge this is the first application of graphical chains models in mental health research and, hence, the first study that assesses the conditional relationships between a multidimensional set of outcome indicators and their predictors.

The three predictors, global functioning, psychopathology and social disability, have been shown to be highly related, thus confirming findings obtained in other studies (Paykel *et al*. 1978; Weissman *et al*. 1978; Serban & Gidynski, 1979; Pai & Kapur, 1982). The baseline value in each variable is able to predict the corresponding value at the 6-month follow-up. But in the subgroup of patients who have a higher severity at baseline, our data point to that patients with a higher severity at baseline improved more than those that had a lower severity.

As expected, higher costs were predicted by lower global functioning and higher psychopathology at baseline. Total direct cost predicted in turn global functioning at 6 month follow-up: higher additional costs were spent for those patients who resulted to be functioning at a lower level 6 month later. Total costs may be considered an indicator of adequacy of the process of care. If resources are correctly assigned, the more severe the patient status at
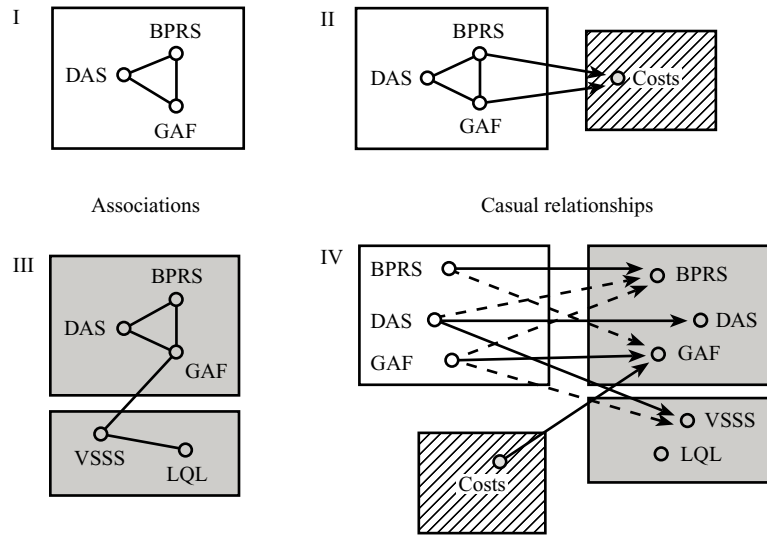
FIG. 1. Graphs of the multivariate dependencies among variables: (I) associations among the set 1 variables; (II) predictors of total costs (set 2 variable); (III) associations among response variables (set 3); (IV) predictors of the responses.

In I and III, the continuous and the dashed line respectively denote direct and inverse correlation. In II and IV, the continuous and the dashed arrow respectively denote direct and inverse direction of the prediction. In order to have the same polarity as BPRS and DAS (low score = good condition, high score = bad condition), in this figure higher GAF scores express higher dysfunction, and higher VSSS and LQL scores express higher dissatisfaction, respectively, with services and with life. To this purpose, in the graph the signs of the associations of the other variables with GAF, VSSS and LQL reported in Tables 2 to 5 were inverted. (□, Wave A; ▨, interval wave–wave B; ▦, wave B.)

baseline the more resources should be prospectively allocated; during the follow-up, the more severe the patient status the more resources should have been spent. Our data show that clinicians in South-Verona behave exactly in this way and direct most effort at those who are most ill in a continuing self-adjusting process.

The outcome indicators appeared to polarize in two groups: an 'objective' disease orientated dimension given by global functioning, psychopathology and disability as assessed by the key-professional, and a 'subjective' disease orientated dimension given by satisfaction with services and quality of life, as assessed by the patients.

Satisfaction with services was predicted by social disability: the less severe patients are, the more satisfied they will be. A puzzling relationship has been found between functioning and satisfaction. At wave B the two variables are correlated: higher functioning corresponds to higher satisfaction with services. However, a low functioning in wave A is a predictor for high satisfaction in wave B. The latter result indicates that an improvement in functioning is correlated to high satisfaction with services. These results

taken together help to clarify some aspects of the ongoing debate on the determinants of satisfaction (Linder-Peltz, 1982) and specifically about the role of expectations, intrinsic characteristics of the patients (such as sociodemographic characteristics, illness severity and so on), and perceived improvement in determining patient satisfaction. From our data we can hypothesize that these determinants may have a different role in different circumstances: the less severely disabled patients at baseline, who probably are also the less demanding, are more likely to be satisfied with services at follow-up; but also those patients with a severe dysfunction at baseline, if they improve after 6 months, are satisfied with services, thus pointing at a powerful role of perceived improvement in determining satisfaction.

Relationships of quality of life with all other variables are mediated by satisfaction with services. A plausible interpretation of this finding could be that both satisfaction with services and quality of life reflect a subjective perception that may be based on a common attitude of the patient. Another possibility is that in psychiatric patients quality of life is strongly determined by

satisfaction with the care they received. Before drawing a definite conclusion further studies are required, since the strong association with satisfaction with services and the relatively small number of patients who took part in our study make it difficult to disentangle the associations pertaining to these two variables; therefore, it could be more prudent to treat them as a dyad.

The sample of this study may be considered representative of patients attending the South-Verona CMHS in the study periods. These patients, in turn, have been shown to be representative of all South-Verona patients receiving psychiatric care (Tansella, 1991). The exclusion of patients who had very few contacts with our CMHS in the period should not have biased our results, considering that these patients are low, and sometimes occasional, users of services. The percentage of patients assessed by the professionals was high. The bias regarding the higher severity of dysfunction of patients refusing to complete the quality of life and the satisfaction with services assessment is more relevant but, unfortunately, is unavoidable in this kind of naturalistic study.

This study has some limitations. The first is that, because of the relatively small size of our sample, we had to reduce the number of variables included in the model. Thus, the potential of the graphical chain model in exploring a very high number of relationships between variables, eliminating redundancies and determining which are the fundamental associations, could not be fully exploited. Specifically, it would have been interesting to explore the role of each instrument's dimensions in the model. With the exception of GAF, in fact, all instruments used have a high number of items, which may be grouped into various dimensions. In some cases the heterogeneity of the various dimensions' content is high and this casts some doubts on the appropriateness of using the instrument overall score. LQL, for instance, is an example: areas like religion, finances, leisure time, health and self-concept are collapsed together in an overall score. With a larger sample it would be useful to run the model after entering into the graphical chain all dimensions scores of each instrument as well as a higher number of predictive variables, including sociodemographic characteristics, psychiatric history and detailed information on service utilization.

Another consideration regards the temporal assessment. In this study, GAF, BPRS and DAS scores have been considered as both predictors and indicators of outcome, and thus they have been assessed twice, at a 6-month interval. This allowed us to take into consideration not only their absolute values but also their change. This was not possible for quality of life and satisfaction with services, which have been assessed only at the end of follow-up. The procedure of assessing only once a year quality of life and satisfaction with services is to be considered legitimate and is widely used (Lehman, 1983) but some authors have stated that the change in the ratings made on separate occasions is a better indicator of outcome (Skanze *et al.* 1992).

We planned to study outcome at 6 months. However, for practical reasons, it was not possible to assess patients at fixed time-intervals. Still, individual total costs were calculated for the actual length of the $time_1$–$time_2$ interval that occurred in each patient. Since about 70 % of subjects have time intervals between 4 and 6 months we expect that the bias introduced by neglecting fixed time assessments will be minor, while this procedure has the advantage to link very precisely the intervention provided with its outcome. Finally, the same model needs to be applied to predictive variables and to outcome variables assessed at a longer follow-up interval: it may in fact be hypothesized that a 6-month period is too short for detecting major changes in the condition of psychiatric patients or major effects of treatments.

In conclusion, graphical chain models have been demonstrated to be an appropriate methodology to analyse process and outcome data. This study throws some light on the complex relationships between clinical, social and economic variables affecting the outcome of mental health care. Specifically, our data show that, in the short-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; those who are more ill get most support. These findings are useful in order to formulate hypotheses on the outcome of psychiatric care that should be tested in future studies.

REFERENCES

Amaddeo, F., Beecham, J., Bonizzato, P., Fenjo, A., Knapp, M. J. & Tansella, M. (1997*a*). The use of a case register to evaluate the costs of psychiatric care. *Acta Psychiatrica Scandinavica* **95**, 189–198.

Amaddeo, F., Bonizzato, P. & Tansella, M. (1997*b*). Psychiatric case registers for monitoring service utilization and evaluating its costs. In *Making Rational Mental Health Services* (ed. M. Tansella), pp. 177–198. Il Pensiero Scientifico Editore: Roma.

Amaddeo, F., Beecham, J., Bonizzato, P., Fenyo, A., Tansella, M. & Knapp, M. (1998). The costs of community-based psychiatric care for first-ever patients. A case register study. *Psychological Medicine* **28**, 173–183.

Attkisson, C., Cook, J., Karno, M., Lehman, A., McGlashan, H., Meltzer, H. J., O'Connor, M., Richardson, T., Rosenblatt, A., Wells, K., Williams, J. & Hohmann, A. A. (1992). Clinical services research. *Schizophrenia Bulletin* **18**, 561–626.

Beecham, J. (1995). Collecting and estimating costs. In *The Economic Evaluation of Mental Health Care* (ed. M. Knapp), pp. 61–82. Arena: Aldershot.

Biggeri, A., Rucci, P., Ruggeri, M. & Tansella, M. (1996). Multidimensional assessment of outcome. The analysis of conditional independence as an integrated tool to model the relationships between variables. In *Mental Health Outcome Measures* (ed. G. Thornicroft and M. Tansella), pp. 207–216. Springer-Verlag: Berlin.

Cox, D. R. & Wermuth, N. (1994). Tests of linearity, multivariate normality and the adequacy of linear score. *Applied Statistics* **43**, 347–355.

Cox, D. R. & Wermuth, N. (1996). *Multivariate Dependencies. Models, Analysis and Interpretation*. Chapman & Hall: London.

Edwards, D. (1993). *Graphical Modelling with MIM 2.2*. Statistical Research Unit, University of Copenhagen.

Edwards, D. (1995). *Introduction to Graphical Modelling*. Springer-Verlag: New York.

Endicott, J. & Spitzer, R. L. (1976) The Global Assessment Scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry* **33**, 766–771.

Everitt, B. S. (1984) *An Introduction to Latent Variable Models*. Chapman and Hall: London.

Fahrmeir, L. & Tutz, G. (1994) *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer-Verlag: New York.

Jenkins, R. (1990). Toward a system of outcome indicators for mental health care. *British Journal of Psychiatry* **157**, 500–514.

Jöreskog, K. G. (1981). Analysis of covariance structure. *Scandinavian Journal of Statistics* **8**, 65–92.

Knudsen, H. C. & Thornicroft, M. (1996). *Mental Health Service Evaluation*. Cambridge University Press: Cambridge.

Kreiner, S. (1996). An informal introduction to graphical modelling. In *Mental Health Service Evaluation* (ed. H. C. Knudsen and G. Thornicroft), pp. 156–175. Cambridge University Press: Cambridge.

Lauritzen, S. L. & Wermuth, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *Annals of Statistics* **17**, 31–54.

Lehman, A. F. (1983). The well being of chronic mental patients: assessing their quality of life. *Archives of General Psychiatry* **40**, 369–373.

Linder-Pelz, S. (1982). Social psychological determinants of patient satisfaction: a test of five hypotheses. *Social Science and Medicine* **16**, 583–589.

Lukoff, D., Neuchterlein, K. & Ventura, J. (1986). Manual for expanded Brief Psychiatric Rating Scale (BPRS). *Schizophrenia Bulletin* **4**, 594–602.

Mirin, S. M. & Namerow, M. J. (1991). Why study treatment outcome? *Hospital and Community Psychiatry* **42**, 1007–1013.

Oliver, J. P. (1991). The social care directive: development of a quality of life profile for use in community services for the mentally ill. *Social Work and Social Science Review* **3**, 4–45.

Pai, S. & Kapur, R. L. (1982). Impact of treatment intervention on the relationship between dimensions of clinical psychopathology, social dysfunction and burden on the family of psychiatric patients. *Psychological Medicine* **12**, 651–658.

Paykel, E. S., Weissman, M. M. & Prusoff, B. A. (1978). Social maladjustment and severity of depression. *Comprehensive Psychiatry* **19**, 121–128.

Roverato, A. & Whittaker, J. (1993). *Standard Errors for the Parameters of Graphical Gaussian Models*. Technical Report. Department of Mathematics. University of Lancaster (UK).

Ruggeri, M. & Dall'Agnola, R. (1993). 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. *Psychological Medicine* **23**, 511–523.

Ruggeri, M. & Tansella, M. (1995). Evaluating outcome in mental health care. *Current Opinion in Psychiatry* **8**, 116–121.

Ruggeri, M. & Tansella, M. (1996). Individual patient outcomes. In *Mental Health Service Evaluation* (ed. H. C. Knudsen and G. Thornicroft), pp. 281–295. Cambridge University Press: Cambridge.

Ruggeri, M., Dall'Agnola, R., Agostini, C. & Bisoffi, G. (1994). Acceptability, sensitivity and content validity of VECS and VSSS in measuring expectations and satisfaction in psychiatric patients and their relatives. *Social Psychiatry and Psychiatric Epidemiology* **29**, 265–276.

Ruggeri, M., Dall'Agnola, R., Greenfeld, T. & Bisoffi, G. (1996). Factor analysis of the Verona Service Satisfaction Scale-82 and development of reduced versions. *International Journal of Methods in Psychiatric Research* **6**, 23–38.

Schulberg, H. C. & Bromet, E. (1981). Strategies for evaluating the outcome of community services for the chronically mentally ill. *American Journal of Psychiatry* **138**, 930–935.

Serban, A. & Gidynski, C. B. (1979). Relationship between cognitive defect, affect response and community adjustment in chronic schizophrenics. *British Journal of Psychiatry* **134**, 602–608.

Skantze, K, Malm, U., Denker, S. J., May, P. R. A. & Corrigan, P. (1992). Comparison of quality of life with standard of living in schizophrenic outpatients. *British Journal of Psychiatry* **74**, 1217–1219.

SPSS Inc. (1996). *SPSS 7.0. User's Guide*. SPSS Inc.: Chicago.

Sytema, S., Micciolo, R. & Tansella, M. (1997). Continuity of care for patients with schizophrenia and related disorders: a comparative South-Verona and Groningen case-register study. *Psychological Medicine* **27**, 1355–1362.

Tansella, M. (ed.) (1991). *Community-Based Psychiatry. Long-Term Patterns of Care in South-Verona. Psychological Medicine* Monograph Supplement 19, Cambridge University Press: Cambridge.

Tansella, M. & Ruggeri, M. (1996). Monitoring and evaluating a community-based mental health service: the epidemiological

approach. In *The Scientific Basis of Health Services* (ed. R. Smith and M. Peckham), pp. 160–169. British Medical Journal Publishing Group: London.

Ventura, J., Green, M., Shaner, 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.

Weissman, M. M., Prusoff, B. A. & Thompson, W. D. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *Journal of Nervous and Mental Disease* **166**, 317–326.

Wermuth, N. & Lauritzen, S. L. (1990). On substantive research hypotheses, conditional independence graphs and graphical chain models. *Journal of the Royal Statistical Society*, *Series B* **52**, 21–50.

Whittaker, J. (1990). *Graphical Models*. Wiley: Chichester.

World Health Organization (1988). *Disability Assessment Schedule (DAS-II)*. WHO: Geneva.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidedlines*. WHO: Geneva.

Wright, R. G., Heiman, J. R., Shupe, J. & Olvera, G. (1989). Defining and measuring stabilization of patients during 4 years of intensive community support. *American Journal of Psychiatry* **146**, 1293–1298.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of American Statistical Association* **57**, 348–368.