A comparative analysis of graphical interaction and logistic regression modelling: self–care and coping with a chronic illness in later life

By Vanessa Didelez

Department of Statistical Science

University College London

United Kingdom

Iris Pigeot

University of Bremen

Germany

Kathryn Dean and Andrew Wister

Gerontology Research Centre

Simon Fraser University

Canada

Summary

Quantitative research especially in the social, but also in the biological sciences has been limited by the availability and applicability of analytic techniques that elaborate interactions among behaviours, treatment effects, and mediating variables. This gap has been filled by a newly developed statistical technique, known as graphical interaction modelling. The merit of graphical models for analyzing highly structured data is explored in this paper by an empirical study on coping with a chronic condition as a function of interrelationships between three sets of factors. These include background factors, illness context factors, and four self–care practices. Based on a graphical chain model, the direct and indirect dependencies are revealed and discussed in comparison to the results obtained from a simple logistic regression model ignoring possible interaction effects. Both techniques are introduced from a more tutorial point of view instead of going far into technical details.

1 Introduction

Choosing appropriate statistical models for data analysis is a key aspect of obtaining valid and meaningful results in investigations of population health issues. Alternative models available for analyzing multivariate data have different strengths and limitations for providing information about how variables affect outcomes of interest. The purpose of this paper is to illustrate what can be learned from comparing and contrasting different statistical models. Two approaches to data analysis, logistic regression and graphical interaction models, will be compared by giving an overview of their foundations and illustrating the information they provide using an empirical example. While logistic regression is well–known, graphical models are more recently developed statistical models with the special feature that certain independence properties among the variables under consideration can be read off a graph. Moreover, we consider especially chain models that focus on revealing indirect associations and identifying hidden and partial relationships. We use an empircal study for illustration but the main emphasis is on the specific theoretical properties of these chain models as opposed to a simple regression.

The paper is organized as follows. After describing the empirical study, i.e. the underlying sample and the variables under investigation as well as the methods used for their measurement, Section 3 gives an introduction to the statistical models which will be applied for analyzing the data set. We start with a brief survey of the theoretical background of a logistic regression and then summarize the basic ideas of graphical models before entering into the discussion of graphical chain models in more detail. One of the most challenging phases of a data analysis is presumably that of selecting an appropriate model for the data. We therefore present the selection strategy used for our analysis where some of the problems which occurred during its application are already briefly addressed. The results obtained from both methods — logistic regression and graphical interaction analysis — are given and compared in Section 4. The final discussion picks up the most important aspects related to our analysis and gives a summarizing assessment of the results.

2 Case Example: Coping with a Chronic Illness

Population aging, coupled with the transition to the chronic disease era, has called attention to the importance of facilitating health maintenance among persons with chronic diseases (WISTER, 1995). The field of self–care and aging specifically addresses this issue. It is now well documented that the majority of illness-related care is self–care, and that health maintenance depends on effective self–care (DEAN, 1989, 1992; ORY and DEFRIESE, 1998). A growing area of interest is the impact of different self–care strategies. Research indicates that self-treatment of disease tends to be helpful in reducing symptoms and the duration of illness episodes. In fact, a large number of studies have shown that self management of disease, especially chronic conditions, has a positive inpact on both physical and psychological health. However, research in this field has not adequately documented the relative importance of different self–care practices and how they interact with other variables to affect one's ability to cope with chronic illness.

The subsequent analyses examine coping with a chronic condition as a function of three sets of variables. These include: 1) background factors (age and sex); 2) illness context factors (the type of chronic illness, duration, and perceived seriousness); and 3) four self–care practices representing diverse strategies of self-management of chronic illness (exercise, reduction of stress, mutual aid, and prescription drug use).

The data have been collected as a part of the baseline of the North Shore Self–Care Study, conducted in 1995 and 1996. Random digit dialling was used to identify prospective respondents meeting the study criteria. Trained interviewers completed telephone surveys with 904 older adults aged 50 and over living in North and West Vancouver communities in Canada. In order to be included in the study, respondents had to have been professionally diagnosed with one of four major chronic illnesses, including: arthritis or rheumatism, heart problems, hypertension, and stroke. A screening question was used to classify the chronic illness perceived to be most serious by the respondent. The current study omits the small number (25) of persons with stroke, resulting in a subsample of 879 persons.

The sample of 879 persons included 417 (47.4%) persons with arthritis/rheumatism, 229 (26.1%) with a heart problem, and 233 (26.5%) with hypertension. The average age of the sample was 69.3 with a range of 50 to 94. Of these persons, 513 (58.4%)

were females and 366 (41.6%) were males.

The dichotomous response variable selected for both sets of analyses is based on the question: "Overall, how well do you feel that you are coping with your condition?" The frequencies for this question show a highly skewed distribution: 11 (1.3%) not at all successful, 44 (5%) a little successful, 368 (42.2%) moderately successful, and 450 (51.5%) very successful (6 missing). We dichotomized this response set into: 0 = persons reporting less than very successful coping (423, 48.5%); and 1 = those reporting very successful coping (450, 51.5%). The rationale for this measurement decision is that the gerontology literature clearly demonstrates that older adults tend to over–rate their well being, as well as subjective dimensions measuring health status, making it a common practice to split skewed distributions in this way.

Nine variables were chosen for analysis. Quantitative variables (i.e. age, time since diagnosis) were recoded into discrete measures. Age at the time of survey is measured using the following categories: 50 - 64, 65 - 74 and 75+, where these categories were chosen to achieve nearly balanced sample sizes in each group. Comparing age groups that have been classified as "young old", "old", and "old old" is considered not only of interest but important in research on aging, where it is common sense that there seems to be no linear age differences in coping (cf. ORY, ABELES and LIPMAN, 1992). A similar rationale is behind categorizing time since diagnosis which was also asked of respondents. Here, it is anticipated that coping with a chronic condition is qualitatively different for recently diagnosed peopole compared to those who have had the condition for some time and those who have managed the condition for many years. Thus, time since diagnosis was recoded into: less than 5 years (275; 31.3%), 5 to 15 years (339; 38.6%) and greater than 15 years (265; 30.1%). Perceived seriousness of the chronic condition was measured using responses to the question: "How serious is your condition?" The response set included: not at all serious (199; 22.6%), slightly serious (254; 28.9%), moderately serious (352; 40%), and extremely serious (74; 8.4%).

Several coping behaviours were measured using yes/no responses to the following question predicated by the opening statement: "People try various things to improve or cope with their condition." ... "Which of these are you doing now?" The coping strategies included in the present study: 1) exercise or become physically active (607; 69%), 2) consult anyone who has the same condition as you (mutual aid, 213;

24.2%), 3) try to reduce stress (426; 48.5%), and 4) take medications as prescribed by your doctor (611; 69.5%).

3 Analytic Techniques

Logistic regression and graphical interaction analyses are applied independently to the data so that similarities and differences can be illustrated on the same substantial background. The theoretical background of the logistic regression is first presented, followed by the graphical interaction analysis. Since the aim of the paper is to compare both techniques we will describe the necessary details but without going too far into formal aspects.

3.1 Logistic Regression

Logistic regression, a widely used statistical model, is a technique developed to assess the influence of explanatory variables (or covariates) on a dichotomous response. Since binary outcomes are often used to describe success/failure situations, logistic regression is one of the most common statistical techniques in health and social sciences. The covariates may be continuous or categorial but since the case of continuous ones is well known and many if not most variables in applications in the health and social sciences are categorical or ordinal measures, we restrict our explanations to this latter situation.

To illustrate the model assumptions, consider a dichotomous response Y, as in our example $Y = 1$ for very successful coping, $Y = 0$ for less than very successful coping, and a couple of explanatory variables, $X_1, \ldots, X_R$, where X_r has I_r categories or different levels which are denoted by x_r. For example X_1 could be the age at the time of the survey with the three ordinal levels as explained above. Then, the main effects logistic regression model states that the probability $\pi(x_1, \ldots, x_R)$ for a success is given as a logistic function of the explanatory variables which may equivalently be formulated as

$$\log \left(\frac{\pi(x_1, \ldots, x_R)}{1 - \pi(x_1, \ldots, x_R)} \right) = \beta_0 + \sum \beta_{x_r}^r, \tag{1}$$

where $\beta_{x_r}^r$ is the effect of the r–th explanatory variable when being in category x_r and β_0 a constant. This implies that the logit of $\pi(x_1, \ldots, x_R)$ is a linear function of the betas. The logistic regression model thus assumes an effect of each level of

each covariate. Note that this is different from the case of continuous covariates where one can hardly speak of effects of specific levels. Since the above model is overparameterized we follow the convention that $\beta_{x_0}^r = 0$ for a reference category x_0 of X_r. In the ordinal case this is mostly chosen to be the highest category. Then, the parameters of the logistic regression are most suitably interpreted in terms of odds ratios. The simple exponential transformation of the beta coefficient $\exp(\beta_{x_r}^r)$ yields the odds ratio for persons with category x_r of X_r as compared to the reference category and with all other explanatory variables unchanged. If the beta coefficients of a specific covariate equal zero for every non–reference category then the probability of success, $\pi(x_1, \ldots, x_R)$, is independent of this covariate given the remaining covariates and given that the main effect model is appropriate. Since the distribution of the response is fully specified by the probability for success it further follows that if the beta coefficients of a specific covariate equal zero for every non–reference category then the response and this covariate are conditionally independent given the remaining covariates, e.g. if X_1 is this covariate we have that $P(Y = 1|X_1, \ldots, X_R) = P(Y = 1|X_2, \ldots, X_R)$. This is the basic idea of variable selection in regression models and an important link to graphical models where the aim is to represent conditional independencies as discussed in the following section (WHITTAKER, 1990).

In this context, it is also important to note that the above model is a main effects model and that the interpretation in terms of odds ratios and possible conditional independence is tied to this model assumption. It is a restriction which may be imposed by the researcher and which has to be justified from the substantial background. However, in many practical situations it represents a serious limitation. Consider for example in our application the covariates age (A) and exercise (E). One could have reasons to assume that the effect of exercise on successful coping is not the same for the different age–levels, e.g. exercise may be more important for successful coping for older persons than for younger ones. If this is true, this means that the effect of one explanatory variable is not the same for different levels of another explanatory variable. This is precisely the meaning of a first order interaction which has to be included in the above model if it is supposed that it has to be accounted for. Since these and even higher order interactions are not allowed by the main effects model it constitutes the most restrictive model in this class. In contrast, a logistic regression model that allows all possible interaction effects of the categorial covariates on the response is always correct, i.e. it contains no restrictions at all. An adequate modelling thus combines statistical evidence and background

knowledge to carefully select the interaction terms. However, for technical reasons
we restrict our application to the main effects and the first order interaction model:
With a sample size of $n = 879$ and the available standard software it is not possible
to carry out any sensible selection in the framework of a logistic regression. Since
the logistic regression approach is thus bound to be unsatisfactory, in the present
data situation, we resort to graphical chain models and alternative methods of in-
cluding and selecting interactions. However, we should bear in mind that it is in
principle possible to include interactions in the logisitic regression model and that
it is of course advisable to do so when interactions are anticipated.

3.2   Graphical Models

Graphical models are a broad class of statistical models for multivariate random
observations whose independence structure can be portrayed in a conditional inde-
pendence graph. In contrast to regression models, graphical modelling is concerned
with identifying association structures for all study variables, including those which
usually are regarded as explanatory. They are therefore appropriate in situations
where complex associations have to be dealt with as in the present study. Due to
the visualization in graphs, these models make it easier to display complex depen-
dence structures. Furthermore, the restriction to the class of graphical models is
often accompanied by technical simplifications of estimating and testing procedures.

In the presentation of a graphical model the vertices stand for the variables and the
edges for associations, where directed edges, i.e. arrows, point from an explanatory
to a response variable and thus indicate asymmetric associations. More precisely, it
is the absence of an edge which is meaningful in the sense that unconnected variables
are conditionally independent given the remaining variables which are prior to or
at the same level as the considered ones: Given a graph $G = (V, E)$ where V is the
set of vertices representing the components of a multivariate random vector $X_V = (X_1, \ldots, X_R)$ and $E \subseteq V \times V$ is a set of ordered pairs such that $(a, b) \wedge (b, a) \in E$
corresponds to a symmetric association and an undirected edge in the graph whereas
$(a, b) \in E \wedge (b, a) \notin E$ corresponds to an arrow, then the basic assumption reads as
follows:

$$X_a \perp\!\!\!\perp X_b | X_{V^* \setminus \{a,b\}} \text{ whenever } (a, b) \text{ and } (b, a) \notin E, \tag{2}$$

where V^* consists of all variables prior to or at the same level as a and b and the
symbol $X \perp\!\!\!\perp Y | Z$ stands for conditional independence between X and Y given Z.

7

Note that the model assumptions are solely concerned with conditional independencies of pairs of variables given the prior variables, without any further assumptions concerning a more simple parametric specification of the probabilities. Therefore, the above conditional independencies represent the key concept of graphical models. Additionally, they lead to a factorization of the multivariate density and thus to a decomposition into 'smaller' models. It can be shown that the latter are essentially given by the so–called cliques of the graph which are maximal complete subgraphs in the sense that on the one hand all variables within the clique are associated and on the other hand adding one of the remaining variables to this subgraph would lead to a loss of this property. The cliques can easily be read off a given graph and represent the basic modules for the multivariate association structure. Property (2) is called the Markov property of the graph and it is the characteristic feature of graphical models as compared to similar multivariate techniques such as path analysis.

Graphical models were developed by DARROCH, LAURITZEN, and SPEED (1980) in the framework of loglinear models and extended to cope with continuous variables by LAURITZEN and WERMUTH (1989). Graphical chain models have been proposed by WERMUTH and LAURITZEN (1990) and are able to capture the dependence structure in more detail since they take account of substantial knowledge to form the dependence chain. Since graphical modelling may be unfamiliar to most of the readers, we attempt to explain this approach in more detail without dwelling too much on the mathematical background. For introductory monographs see EDWARDS (2000) and WHITTAKER (1990) and for a thorough mathematical treatment of the subject see LAURITZEN (1996).

Graphical chain models

In our application, we consider graphical chain models which are characterized by the assumption of an underlying recursive structure of the variables, i.e. the latter have to be ordered into a sequence of subsets of variables referred to as recursive levels of the model. This means that starting with pure explanatory variables these may influence some responses which again might be influential for others and so on, until we arrive at the pure responses which are possibly influenced by all variables but do not in turn influence any additional variables. This recursive structure has to be fixed a priori by the researcher and represents in most cases assumptions on the data generating process based on the temporal and/or hypothesized causal structure. Clearly, the recursive structure has similarities to the relationship of response and explanatory variables within a regression model. But in contrast to usual re-

gression models, graphical chain models allow for variables that are simultaneously response and explanatory variables — so–called intermediates which account for possibly indirect influences.

Figure 1. Hypothetical graph for age (A), exercise (E), and coping (C).
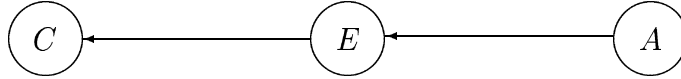
To illustrate this, consider again the variables age, exercise, and coping. A possible assumption on the dependence structure may be that successful coping directly depends on exercise in the sense that physical activity improves coping with the illness. That is, age may not affect coping directly, but instead operate through intervening influences like exercise. Thus, age indirectly affects coping. This dependence structure can be read off the graph in Figure 1. In this simple model, each variable is on a different recursive level. The separation of the vertices C and A by E implies the conditional independence of C and A given E, i.e. $P(C, A|E) = P(C|E)P(A|E)$. The indirect effect of A on C would not be noted in a usual regression model since it is mediated by E. It should be emphasized that the direction of an arrow is not determined by data or statistics but only because of subject matter knowledge.

Usually, a recursive level is formed by more than one variable resulting in a multivariate regression of the variables at this level given those at lower levels. Variables at the same level are assumed to be on equal footing, i.e. no sensible reponse–explanatory relationship can be assumed within this subset. In the graphical representation, these form a so–called chain component. In contrast, dependencies between variables in different components are considered as directed from earlier or lower levels (represented on the right hand side) to subsequent higher levels. The remaining variables inbetween are intermediates. The recursive structure therefore partitions the vector X_V into subvectors $X_{V(1)}, \ldots, X_{V(K)}$ such that the joint density $f(x_V)$ is decomposed into a product of conditional distributions as follows

$$f(x_V) = f(x_{V(1)}) \prod_{i=2}^{K} f(x_{V(i)}|x_{V(1)}, \ldots, x_{V(i-1)}). \tag{3}$$

Each of the factors in (3) corresponds to the distribution of variables at one level conditional on variables at all lower levels. Thus, one may regard a graphical chain

model as a sequence of regression models that describe these conditional distributions and the choice of the recursive structure reflects that one is specifically interested in the latter. In Figure 1, the indirect association of A and C via E would thus be revealed by the additional regression of E on A. If there is only one variable at a given level, the corresponding regression model is a standard multiple regression model with a single dependent variable. If the dependent variable happens to be a binary variable, as in our application the variable coping, it will be a logistic regression model with the only unusual feature that the model includes higher order interaction parameters, as will be explained below.

Note that a factorization like (3) is possible for an arbitrary partition, but in our case they should be meaningful from a substantive point of view as follows: The researcher decides which variables are at the same recursive level, i.e. in the same block, and how to arrange the blocks. Afterwards it is the task of the statistical model selection to detect a finer structure, i.e. to decide on the basis of appropriate tests on conditional independencies which undirected and directed edges have to be added in order to fit the data appropriately. The chosen recursive structure should therefore permit to reveal as much conditional independencies as possible. As noted by LAURITZEN and RICHARDSON (2002), such a structure might not necessarily correspond to the temporal order in case of hidden relevant variables. Let us consider two simple examples in order to clarify the consequences of the choice of the ordering: (i) Let X_1 and X_2 be two explanatory variables which are marginally independent and Y a response variable which depends on the former. If these are arranged in an adverse order, i.e. Y prior to X_1 and X_2, then per definition of a chain graph as given in (2) the relation between X_1 and X_2 will always be examined conditionally on Y and will thus only show that they are conditionally dependent but not that they are marginally independent. (ii) In the second example let X be an explanatory variable and Y_1 and Y_2 responses which are independent of each other when conditioning on X. Again, this conditional independence could not be detected if these variables are arranged in an adverse order, i.e. Y_1 and Y_2 prior to X.

To illustrate the interpretation of a chain graph consider the recursive structure given in Figure 2.

Figure 2. Recursive structure for variables in this study. ($C =$ coping, $E =$ exercise, $S =$ stress reduction, $M =$ mutual aid, $P =$ prescription drug use, $X =$ perceived seriousness, $I =$ chronic condition, $T =$ time since diagnosis, $G =$ gender, $A =$ age).
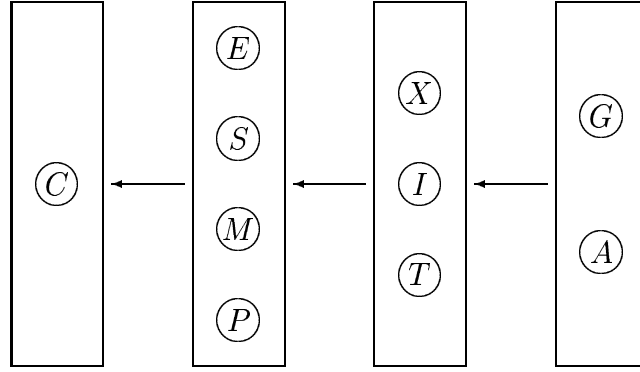
This has to be interpreted as follows. In our application coping with the chronic condition is the only pure response variable and gender as well as age are clearly pure explanatory since they may not be influenced by any other variables. Variables on equal footing are for example the self–care strategies exercise (E), stress reduction (S), mutual aid (M), and taking prescribed medications (P) in the second box. By the recursive structure it follows that taking prescribed medications may be influenced for example by the chronic condition or age but not by coping with the chronic condition. The possibility of including such intermediates in the model constitutes the most important difference between simple regression models as logistic regression and graphical chain models. It offers the opportunity to find indirect associations. In the assumed recursive structure of our study as depicted in Figure 2, age could influence coping via the mediating variables perceived seriousness (X) and exercise (E). But it is also possible that age has a direct effect on coping. This first rough dependence chain is now the basis for finding the final graph by appropriate tests on exactly those conditional independencies given in (2), for instance an edge between X and S is added if these are found to be conditionally dependent given all variables at the same or prior levels, E, M, P, I, T, G, and A. In the following paragraphs, we give the theoretical background for the derivation of appropriate tests.

Graphical loglinear models

The Markov properties (2), i.e. the postulation of conditional independence whenever there is no edge between two variables, impose specific restrictions on the underlying statistical model which have to ensure that the model class reflects exactly those independencies shown in the graph, no less and no more. This is important since any further analysis relies on the validity of the Markov properties. The corresponding restrictions have been explored for different kinds of graphs and different classes of distributions, the most general being the so–called Conditional Gaussian

distribution (WERMUTH and LAURITZEN, 1990). In case of an undirected graph and all variables categorial, graphical models are given as a subclass of loglinear models. We thus restrict our considerations for the moment to the latter, the connection to chain models which is the relevant model class in the present application is given below. In case of a discrete random vector X_V the loglinear model is in the most general case given by the assumption that

$$\log P(X_V = x_V) = \sum_{d \subseteq V} u_d(x_d),$$

where u_d is a constant for $d = \emptyset$, main effects for $|d| = 1$ and interactions otherwise. Graphical loglinear models are a subclass of hierarchical loglinear models. The latter have the property that if an interaction of a certain order is included in the model then the corresponding lower order interactions and main effects have to be included, i.e. if u_d is not restricted to be zero then neither is $u_e, e \subset d$. Hierarchical models can thus be characterized by the sets of variables forming the highest order interactions which is called the generating class. These models are graphical, i.e. they hold the Markov properties (2) of a specific graph, if the generating class is identical with the cliques of a graph (DARROCH et al., 1980). Let $\mathcal{C}$ be the set of cliques then the graphical loglinear model reads as

$$\log P(X_V = x_V) = \log p(x_V) = \sum_{C \in \mathcal{C}} \sum_{d \subseteq C} u_d(x_d).$$

In Figure 1, for instance, the cliques are $\{C, E\}$ and $\{E, A\}$. Thus, the corresponding graphical loglinear model is given by

$$\log p(x_1, x_2, x_3) = u_\emptyset + u_C(x_1) + u_E(x_2) + u_A(x_3) + u_{CE}(x_1, x_2) + u_{EA}(x_2, x_3).$$

The fact that the u_{CA} interaction and the second order interaction u_{CEA} are not included in the model means that the desired conditional independence of C and A given E holds.

Graphical loglinear models can easily be modified to apply to chain graphs, i.e. to model directed dependencies, by conditioning. Consider again the example in Figure 1: If the joint distribution $P(C, E, A)$ follows the given graphical loglinear model then the conditional distribution $P(C|E, A)$ corresponds to a logistic regression with a main effect of E and no effect of A since

$$
\begin{aligned}
\log \left(\frac{\pi_C(x_2, x_3)}{1 - \pi_C(x_2, x_3)} \right) &= \log p(1, x_2, x_3) - \log p(0, x_2, x_3) \\
&= u_C(1) + u_C(0) + u_{CE}(1, x_2) - u_{CE}(0, x_2),
\end{aligned}
$$

12

such that $u_C(1) + u_C(0)$ corresponds to β_0 and $u_{CE}(1, x_2) - u_{CE}(0, x_2)$ to $\beta_{x_2}^2$ in (1). Similarly the conditional distribution $P(E|A)$ corresponds to a multinomial regression with a main effect of A. Recall that we always have to condition on all the variables at lower levels but not on those at higher levels in order to capture the Markov properties of the chain graph. Besides the already mentioned consideration of intermediates an important difference to logistic regression is the inclusion of higher order interactions.

A special feature of graphical loglinear models is that collapsible sets can easily be identified. Collapsibility implies that the marginal distribution of a subset of the variables follows the dependence structure of the corresponding subgraph and thus can easily be read off the graph. This is done automatically by most of the software for graphical models so that we resist from going into details. In Figure 1 for instance, the graph is collapsible onto the set $\{C, E\}$ but not onto $\{C, A\}$ since marginalization with respect to E would introduce a dependence between C and A whereas the subgraph $G_{\{C,A\}}$ is an independence graph. Collapsibility is a very convenient and important property because it ensures that the dependence structure within the corresponding subset of variables can be analyzed without having to take the remaining variables into account, i.e. it helps in identifying sufficient sets of variables relevant for assessing a pairwise conditional relation. Since the relevant variables will typically be a true subset of the remaining variables, collapsibility is the key to reducing computational effort and to increase the power of appropriate tests of pairwise conditional independence.

Model selection in this study

Theoretically, it would be the best strategy to try all potentially adequate graphical models and evaluate their goodness of fit. This is, however, in any realistic application impossible due to the necessarily enormous effort. Instead, one has to face the demanding task to find a selection strategy that calculates a graphical model on the basis of a given data set and which allows for the above described interpretation of the resulting graph in terms of the Markov properties, i.e. all the conditional independencies represented by the graph should be supported by the data. In addition, an adequate strategy should try to yield a parsimonious model being consistent with the data, i.e. the graph should only show as much associations as necessary to fit the data. Such a strategy can be based on e.g. forward and backward selections using criteria such as p–values or Akaike's information criterion. Other proposals have for instance been made by COX and WERMUTH (1996), EDWARDS and HAVRÁNEK

(1985) or SPIRTES, GLYMOUR and SCHEINES (1993).

Whensoever a graphical chain model is to be selected, the recursive levels of the chain graph have to be fixed before starting the statistical analysis. In our application we choose the one already depicted in Figure 2. This recursive structure is based on substantial knowledge on the possible causal relationships among the variables and does not depend on the special data given here.

The selection strategy we propose relies on pairwise tests of conditional independence corresponding to (2) for any two variables which are also known as tests of zero partial association (BIRCH, 1965). If, for example, the hypothesis of conditional independence between C (coping) and E (exercise) given M (mutual aid) and I (illness) is to be tested, i.e. the inclusion or exclusion of an arrow from E to C, this can be done by testing the loglinear model with generators $\{CMI\}$ and $\{EMI\}$ against the saturated model $\{CEMI\}$. Additionally, note that we can reduce our considerations to three–dimensional tables since the conditioning variables may be combined into one categorial variable. In the preceding example, define $Z = M \times I$ and test $C \perp\!\!\!\perp E|Z$. From this it can also be seen that it is not necessary to fit regression models. The recursive structure enters the analysis by placing constraints on the model search which prevent from conditioning on sets of variables that are at higher recursive levels than the considered ones and it determines the direction of edges that cross levels. For our analysis, we use the software program DIGRAM (Discrete GRAphical Modelling) developed by KREINER (1992) where approximations to exact p–values of any appropriate test statistic are calculated instead of using the asymptotic distribution. KREINER (1987) has shown that the latter can be very misleading when used for large sparse tables as it is the case with the data analyzed here. The approximation to the exact p–values is performed as described in KREINER (1987): Consider a three–dimensional table given by variables X, Y, and Z. The restriction $X \perp\!\!\!\perp Y|Z$ implies that the sufficient tables are given by the $X \times Z$– and $Y \times Z$–marginal tables. Thus, a test on conditional independence has to condition on these marginals. The corresponding test procedure is carried out according to the following algorithm: (1) Generate a random sample of tables according to the distribution of the table given these sufficient marginals. This can be done successively for the slices of two–dimensional $X \times Y$ tables, one for each level of Z, resulting in samples from ordinary hypergeometric distributions. (2) Calculate for each generated table the appropriate test statistic. (3) Approximate the level of significance as the proportion of tables which are at least as favourable to the

alternative as the observed table.

The p–values generated in this way are of course subject to random error, but this can be controlled whereas only little is known about the behaviour of p–values based on the asymptotic distributions of the corresponding test. The known results suggest that the latter systematically underestimate the true p–values. The exact p–values calculated here are based on 2000 simulation runs which yield a sufficient level of accuracy for our purposes (cf. KREINER, 1987).

The tests on pairwise conditional independence implemented in DIGRAM are either Pearson's χ^2 which is equivalent to the likelihood–ratio test based on the underlying loglinear model assumption or the partial γ–coefficient suggested by DAVIS (1967), where the latter is only computed if both variables under consideration are ordinal or binary. The application of the γ test goes beyond the loglinear framework but since it uses more information, namely on the ordinal structure of the variables, it yields a more powerful test. The partial γ–coefficient of two variables X, Y given Z is calculated as follows: Let $N^s_{X,Y|Z} = \#\{(x^i < x^j$ and $y^i < y^j)$ or $(x^i > x^j$ and $y^i > y^j)$ while $z^i = z^j, i, j = 1, \ldots, n\}$, i.e. the number of concordant pairs in X and Y for constant Z, and $N^d_{X,Y|Z} = \#\{(x^i < x^j$ and $y^i > y^j)$ or $(x^i > x^j$ and $y^i < y^j)$ while $z^i = z^j, i, j = 1, \ldots, n\}$, i.e. the number of discordant pairs conditional on Z. Then,

$$\gamma_{X,Y|Z} = \frac{N^s_{X,Y|Z} - N^d_{X,Y|Z}}{N^s_{X,Y|Z} + N^d_{X,Y|Z}}.$$

The above coefficient ranges between -1 and 1 where $\gamma_{X,Y|Z} = \pm 1$ stands for a purely positive or negative monotone (not necessarily linear) relation, respectively. Inference is based on the smaller of the p–values of both tests and carried out on a 5% significance level. This raises the sensitivity for detecting possible monotone dependencies between ordinal variables and it has been shown that the partial γ–coefficient often discloses interactions not detectable by the likelihood ratio or χ^2 tests (KREINER, 1987). Note that logistic regression neither accounts for an ordinal measurement scale of the explanatory variables nor does the usual software calculate exact p–values. But especially with our data the exact p–values essentially differ from the asymptotic ones.

Let us now come to the selection procedure itself. It is a modification of the standard stepwise selection for regression models and comparable to the one presented in WHITTAKER (1990, Chapter 8). The first step is backward starting from the saturated model, i.e. the complete graph containing all possible edges. The tests of

pairwise conditional independencies in this first step are carried out conditionally on all remaining variables, and then in the subsequent backward steps, on all relevant variables. The relevant variables for a pairwise relation are at each step determined by exploiting the collapsibility properties of the current graph. In contrast to usual stepwise selections, all non–significant edges are excluded at a time. Forward procedures again test all pairwise independencies and all significant edges are included. Note that due to different conditioning sets chosen at each step accordingly to the current graph, the backward and forward procedures do not cancel each other, but yield different results. These steps continue until no further changes take place. Although there is no obvious reason for this strategy to converge in general, we reach a final graph with our data. In our example, the graph resulting from the analysis is shown in Figure 3. It is important to bear in mind that data sets generally support more than one plausible model, and there could conceivably be a model that fits the data even better. A reason is that the omission of any non–significant edge may mean that there is indeed a conditional independence, or it could mean that a conditional dependence cannot be detected because the sample size is too small.

4   Results

The results of the statistical analysis given here should be interpreted with caution, especially regarding their statistical significance since quite a large number of statistical tests is carried out especially during the selection process without accounting for multiplicity. For the sake of simplicity, effects are, in spite of this objection, referred to as statistically significant if the p–value of the corresponding single statistical test is below 0.05. The p–values are reported as additional exploratory information.

4.1   Logistic Regression

For the reasons described in Section 3.1 we restrict ourselves to a main effect logistic regression model where coping is regressed on the basic variables age and sex, the three illness context factors, and on the four self–care practices. The latter four variables are of special interest because on the one hand they form those potential influences on coping which might be positively influenced by an appropriate intervention strategy. On the other hand, the self–care practices are located closest to the main response variable coping in the pre–formulated dependence chain. Note that the overall fit for the main effects model is quite poor (p=0.0006), as anticipated. It can slightly be improved by adding first order interactions but these emerge as

16

not significant and are therefore discarded. Note, also, that the measure of fit is based on an asymptotic distribution which might not be appropriate due to sparse cell frequencies. The mentioned limitations of the analysis should be kept in mind when interpreting the following results.

Coming now to the results in detail, it turns out that for persons engaging in exercise to cope with a chronic condition the odds to cope very successfully are 1.41 times as large as for those who do not exercise, controlling for all other variables. Note that this control is w.r.t. the main–effects–only assumption which is not equivalent to conditioning on all remaining variables as in the graphical model analysis. It is interesting to note that those who engage in mutual aid or stress reduction, as measured in this data, are slightly less likely to report very successful coping than persons not employing these self–care responses, odds ratios of 0.83 and 0.78, respectively. Surprisingly, prescribed medications are not found to influence coping in this data. These results are rather unexpected and lead to nonintuitive conclusions, which can be hardly explained. Several aspects should, however, be taken into account. The surprising findings concerning stress reduction may be a measurement issue since the corresponding question could have been understood in different ways. Further, stress as well as how it is perceived differ considerably among those who try various things in their own way to reduce stress. Nevertheless, the variable should remain in the analysis because of its importance. In addition, coping effectively with a chronic condition has many social and psychosocial components. It appears that taking medicines may not be considered a major determinant of coping with the conditions studied in this investigation.

Besides these results, the main effect of age is rather low with an odds ratio of 1.47 compared to other potential influences as for instance the two chronic conditions heart problem and hypertension. Here, we observe an odds ratio of 3.03 for persons with heart problems compared to those with arthritis and rheumatism. The odds to cope very successfully are for persons in the hypertensive group about two times greater than those from the arthritis group. Furthermore, for persons reporting that their chronic condition is not serious the odds are about three times (odds ratio of 3.24) greater. A weak association for sex can also be seen from the results with males having a slightly lower probability (odds ratio of 0.86) of reporting that they are coping very successfully than females. Finally, years since diagnosis do not come up with a statistically significant influence. The above results are summarized in Table 1.

4.2 Graphical Chain Model

The graph resulting from the above described selection strategy is shown in Figure 3. Not surprisingly, the overall model fit is very good (deviance=2459 with df=6688). This illustrates the drastical effect of the inclusion of all higher interactions whenever there is no conditional independence. However, one has to bear in mind that many slightly different chain graphs will have an equally good fit so that this measure of fit on its own is not really meaningful.

Direct effects: We find that exercise, perceived seriousness, mutual aid, stress reduction, and age are direct influences on coping, i.e. these effects persist even when conditioning on the remaining variables, respectively. They all exhibit associations in the same direction as found in the logistic regression analysis. Due to the use of the γ– and χ^2–statistics the quantification of the strength of associations is different (cf. Table 2; instead of the χ^2 we report the corrected contingency coefficient K^* which is the normed χ^2 such that $K^* = 0$ if there is no association at all and $K^* = 1$ if there is an extreme association). In addition, neither method supports a direct association between coping with a chronic condition and prescription drug use, gender or time since diagnosis. More precisely, we have by the graphical result that these variables are pairwise conditionally independent given all the remaining variables, respectively, as stated by the Markov property (2). However, while these results parallel those generated using the logistic regression, there is also a discrepancy: The graphical model shows no direct effect of chronic condition on coping. Since the first step of the graphical analysis, i.e. analyzing the first recursive level, is equivalent to a logistic regression including all possible interactions, the different result is due to these interactions and possibly to the use of exact tests and the choice of the selection procedure of course. The local results described below will provide further evidence for the presence of higher order interactions.

Based on other selection strategies such as simple backward selection which we additionally applied to get an idea about the stability of our obtained results, it can be stated that the influence of the first three variables, i.e. exercise, perceived seriousness, and mutual aid, seems to be undoubted. The other direct influences in contrast have to be interpreted with some caution since the corresponding pairwise association heavily depends on the choice of the conditioning set which varies steadily during the selection procedure. Especially the missing direct effect of the type of the chronic condition will have to be considered in more detail. In contrast, the missing edges between prescription, gender, and time on the one hand and coping

on the other hand are again stable results.

Indirect effects: Besides the analysis of direct effects the graphical model results provide some evidence for indirect influences going beyond the findings of the logistic regression. First, all those variables that have a direct effect also affect coping through indirect pathways but we will not comment on this. More interestingly, type of illness, gender, and time since diagnosis, i.e. all the remaining variables except prescription drug use, have an indirect effect on coping with a chronic condition and we will comment on the first two (the corresponding results can again be read off Table 2). In terms of the Markov properties (2), indirect influences without direct association mean that there is a conditional independence between for instance coping and type of illness when conditioning on the remaining variables. But if we do not condition at least on G, A, X, S, and M then there will be an association since this is the minimal separating set. Note that indirect influences are also conveyed through symmetric associations.

The chronic condition is symmetrically associated with perceived seriousness such that the arthritis patients given gender tend to a higher perceived seriousness than the others especially than the hypertension group which through the negative direct effect of perceived seriousness implies that the arthritis group tends to cope less than very successfully. Additionally, the chronic condition has an effect on stress reduction and mutual aid. Inspection of the corresponding conditional tables cannot suggest a simple pattern for the illness/stress reduction association, but it is clear that the arthritis group is the one reporting most mutual aid. Moreover, the graphical model suggests that the strong effect of illness type on coping detected by the logistic regression, i.e. the hypertension and heart problem groups coping more successfully than the arthritis group, can be explained 'away' by conditioning on the variables E, S, M, X, and A. Note that conditioning on an arbitrary subset of $\{E, S, X, A\}$ or $\{M, S, X, A\}$ already yields a significant influence of the type of illness on coping implying that all these variables are relevant.

Gender, in turn, is symmetrically associated with age — women tend to be older than men which is characteristic for this particular subpopulation of elderly persons and which results in women coping less successfully because of their age. Additionally, gender also affects stress reduction, perceived seriousness, and illness type, i.e. females are more likely to engage in stress reduction techniques ($\gamma=0.26$), women tend to report a slightly higher perceived seriousness than men ($\gamma=0.11$), and the chronic condition faced by most of the women is arthritis whereas men mainly have

to cope with heart problems. Taking into account that the indirect effect of illness type on coping is such that the arthritis group is less probable to cope very successfully as explained above, these pathways give reasons for the slightly smaller overall probability of women to cope very succesfully than men (the marginal $G \times C$–table shows a non–significant $\gamma = -0.07$).

Local results: The above findings result from analyzing conditional dependencies. It is only natural to ask how each of these pairwise associations changes for the different categories of the conditioning variables where it can be read off the graph which variables are relevant. Note that only if the associations are the same for the different conditions the assumption of no higher interactions made by the main effect logistic regression would be supported. In this paragraph, we present the local results related to the self–care domain, i.e. to the influence of exercise, stress reduction, and mutual aid on coping. Each pairwise dependence is considered at the different levels of each individual variable out of the relevant conditioning set which contains all those that have a direct effect on coping.

From the results given in Table 3 the most striking one is that the effects of exercise, stress reduction, and mutual aid on coping are extreme under the condition that perceived seriousness takes the value 'not at all serious'. The first effect is positive and the latter two are negative corresponding to the global result.

Additionally, the significant effect of exercise on coping given mutual aid, stress reduction, seriousness, and age is clearly replicated for almost all categories of the conditioning variables but its strength is different depending on the condition although it is generally quite high. Considering the interaction of self–care strategies we note that exercise has a slightly weaker effect when mutual aid is present. Besides these findings, it could of course be supposed that the strength of the effect varies for the different types of chronic conditions. The results of the graphical analysis, however, suggest that this is probably not directly due to the illness but to its relations to the remaining variables.

Considering the effect of stress reduction we can see that besides for the $X = 1$ category there is only one significant result for the exercisers (γ=-0.17). Obviously this is enough for the global test to be significant. Note that especially for the youngest age group and when perceived seriousness is moderately serious the negative effect almost vanishes. Since these results seem to be somewhat unsatisfactory they should be given some more attention in future surveys.

The significant global effect of mutual aid is partially replicated in the different cat-

egories. It is low, for instance, for those who do not exercise and again quite strong for those who do exercise. A similar pattern can be found with respect to the age groups. The different aspects due to exercising are again picked up in the discussion.

Another interesting question which might be addressed using the results of the graphical analysis could be concerned with the factors determining which of the self–care practices are applied. It seems for instance to be clear that exercise is a response to perceived seriousness only, while being independent from the type of the chronic condition. But we restrain ourselves from going into details.

5   Discussion

There has been a growing concern that quantitative data analyses are not keeping pace with conceptual, theoretical, and substantive research problems in a variety of fields in the social and biometrical sciences. A necessary prerequisite for advancing knowledge is the development of statistical methods that are better able to address new generations of research questions. This paper compared and contrasted two statistical approaches to quantitative data analysis, logistic regression and graphical interaction modelling, to show how model assumptions and selection strategies affect findings in research investigations.

The findings illustrate how restricting statistical analyses to simple regression models can limit and distort findings in research studies. It is quite often the case that the effects of an influence are not the same for different levels of one or more other explanatory variables. In health and social research, failure to identify the conditions under which influences affect an outcome can result in serious misapplications in policy or practice. That graphical loglinear models include higher order interactions given by the cliques of the graph constitutes an important difference between the two approaches to statistical analysis discussed in this paper. However, as mentioned earlier, it is possible to include interactions in the logistic regression model. If one is specifically interested in the regression parameters the results from the graphical modelling analysis might be useful to decide on which interactions should be included.

While the example used to compare models can only be considered illustrative, it was shown that graphical chain models can use substantive knowledge to capture

complex dependence structures in data collected to study research issues. The findings illustrate, as pointed out in the methods sections, that in contrast to usual regression models, graphcial interaction models allow the explication of variables that are simultaneously response variables and explanatory variables, to identify relevant indirect influences. Since graphical chain models may be considered as a sequence of regression models, they may be used to expand and refine the information provided by traditional regression models. The methods used in this study have the additional advantage that information on ordinal variables is utilized to yield more powerful tests.

The graphical interaction results portray a complex set of direct relationships between the variables in the model, as well as interactions. One could conclude from the logistic regression results that, other things being equal, persons who exercise generally perceive themselves as coping very well more often than those who do not exercise. If we turn to the graphical interaction results it can be seen that the effect is not equally high for the different subgroups formed especially by the categories of perceived seriousness and age. Additionally, we have to conclude that the strength of the effect of exercise on coping does not depend on the type of chronic condition given the other relevant variables. This is a somewhat striking result which may require additional investigation to further disentangle. It might for instance be interesting to identify the specific attributes of a chronic condition that are most relevant for succesfully coping. Such findings and conclusions have clear and profound implications for our understanding of the complexity of the self–care/illness nexus, as well as for health promotion design, implementation, and evaluation. For example, they may be very useful in identifying persons who may be most likely to be at a stage to change their exercise self–care.

Finally, we should balance our discussion with some of the limitations of graphical modelling. As with any quantitative technique, oscillation, refinement, and replication of the theory–research relationship is fundamental to providing the rationale for inclusion and modelling of variables. In addition, since multidimensional contingency table analysis disaggregates data into small sub–sets to which statistical tests are applied, issues of sample size and statistical power need to be addressed. In the present research, we used a sample size of 879, and nine explanatory variables with discrete categories ranging between two and four. Small or zero cell sizes were therefore averted so that although the presence of an edge is ensured by the 5%

level (not taking multiplicity into account), the absence is not necessarily based on strong evidence for conditional independence. Here, statistical power is enhanced by analyzing relationship in cliques and by using exact tests of statistical significance as well as knowledge about ordinal scales. Another general caution has to be formulated which is not restricted to graphical models but applies to all model search techniques. The data generally support several maybe substantially different models so that the model choice heavily depends on the selection strategy. Comparison of the results from different strategies might reveal which relations are the most delicate as to enhance further and more specialized research.

In conclusion, this comparative case study demonstrates several advantages and disadvantages of conventional analytic techniques and graphical interaction analysis of multidimensional contingency tables. While the newly developed techniques respond to several deficiencies in their precursors, it is not a statistical panacea. Graphical modelling does, however, move us in another direction — from simplification of complex human behaviour to reflection of its inherent multidimensional nature. It clearly responds to the need for methodological developments that can better capture linkages between dynamic systems (SMITH and TORREY, 1996; SUSSER and SUSSER, 1996a; 1996b). The identification of antecedent and intervening variables, coupled with elaboration of interactions between variables, makes an important contribution to knowledge development. It is hoped that this paper prompts further applications and evaluations of graphical models in response to the current debates that preoccuppy the social and biological sciences.

Table 1. Results of the main effects logistic regression (odds ratios are w.r.t. the non listed category).

Explanatory variable		O.R.	p–value	Explanatory variable		O.R.	p–value
Age group	50 − 64	1.47	0.001	Seriousness	Not at all	3.24	0.000
	65 − 74	0.92	0.452		Slightly	0.95	0.727
Gender	Male	0.86	0.052		Moderately	0.59	0.000
Illness	Heart prob.	3.03	0.000	Exercise	Yes	1.41	0.000
	Blood press.	2.01	0.000	Mutual aid	Yes	0.83	0.038
Time	< 5	0.91	0.394	Stress reduc.	Yes	0.78	0.002
	5-15	0.92	0.457	Prescribed med.	Yes	0.86	0.083

Table 2. Conditional γ– and K^*–values (the latter in italics) for all non vanishing pairwise relations in the graph (K^* is the normed χ^2 such that $K^* = 0$ if there is no association at all and $K^* = 1$ if there is an extreme association; it is computed for nominal variables).

Effects on coping			Effects on self–care strategies		
Hypothesis	γ / K^*	p–value	Hypothesis	γ / χ^2	p–value
$C \perp\!\!\!\perp E \vert SMXA$	0.31	0.001	$E \perp\!\!\!\perp S \vert X$	0.41	0.000
$C \perp\!\!\!\perp S \vert EMXA$	-0.16	0.036	$E \perp\!\!\!\perp X \vert S$	-0.23	0.000
$C \perp\!\!\!\perp M \vert ESXA$	-0.22	0.009	$S \perp\!\!\!\perp X \vert EIGA$	0.22	0.000
$C \perp\!\!\!\perp X \vert ESMA$	-0.50	0.000	$S \perp\!\!\!\perp I \vert MXGA$	*0.46*	0.034
$C \perp\!\!\!\perp A \vert ESMX$	-0.18	0.002	$S \perp\!\!\!\perp G \vert IXA$	0.26	0.001
Effects on illness context variables			$S \perp\!\!\!\perp A \vert MIXG$	-0.28	0.001
$X \perp\!\!\!\perp I \vert G$	*0.23*	0.001	$M \perp\!\!\!\perp S \vert IA$	0.38	0.000
$X \perp\!\!\!\perp G \vert I$	0.11	0.029	$M \perp\!\!\!\perp I \vert SA$	*0.23*	0.022
$I \perp\!\!\!\perp G \vert XA$	*0.45*	0.000	$M \perp\!\!\!\perp A \vert SI$	**0.24*	0.014
$I \perp\!\!\!\perp A \vert TG$	*0.34*	0.000	$P \perp\!\!\!\perp X \vert IT$	0.35	0.000
$T \perp\!\!\!\perp I \vert A$	*0.23*	0.001	$P \perp\!\!\!\perp I \vert XT$	*0.52*	0.000
$T \perp\!\!\!\perp A \vert I$	0.22	0.000	$P \perp\!\!\!\perp T \vert IX$	0.14	0.026

* This is the only case where the χ^2–test is significant but not the γ, i.e. there is a significant non–monotone association.

Table 3. Local results for the direct effects of self–care strategies on coping.

Level		$C \perp\!\!\!\perp E\|MSXA$		$C \perp\!\!\!\perp S\|EMXA$		$C \perp\!\!\!\perp M\|ESXA$	
		γ	p–value	γ	p–value	γ	p–value
Exercise	No	—	—	-0.11	0.246	-0.15	0.229
	Yes	—	—	-0.17	0.049	-0.24	0.014
Mutual aid	No	0.31	0.001	-0.14	0.067	—	—
	Yes	0.27	0.102	-0.26	0.069	—	—
Stress reduc.	No	0.30	0.004	—	—	-0.27	0.040
	Yes	0.32	0.012	—	—	-0.20	0.054
Seriousness	Not at all	0.59	0.004	-0.49	0.018	-0.49	0.012
	Slightly	0.23	0.056	-0.16	0.136	-0.28	0.051
	Moderately	0.29	0.010	-0.09	0.247	-0.13	0.166
	Extremely	0.24	0.230	-0.26	0.182	-0.25	0.242
Age	50 - 64	0.28	0.036	-0.07	0.335	-0.01	0.459
	65 - 74	0.42	0.003	-0.24	0.063	-0.34	0.009
	≥ 75	0.23	0.061	-0.16	0.151	-0.40	0.019
Global effect		0.31	0.001	-0.16	0.036	-0.22	0.009

Figure 3. Final graph resulting from stepwise selection (C = coping, E = exercise, S = stress reduction, M = mutual aid, P = prescription drug use, X = perceived seriousness, I = chronic condition, T = time since diagnosis, G = gender, A = age).
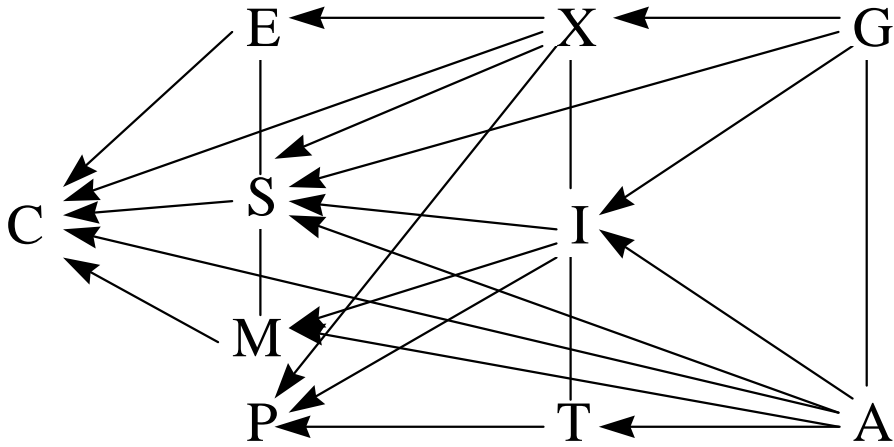
References

BIRCH, M.W., 1965: The detection of partial association, II: The general case. *Journal of the Royal Statistical Society, Ser. B*, **27**, 111–124.

COX, D.R. and WERMUTH, N., 1996: *Multivariate Dependencies: Models, Analysis and Interpretation.* Chapman and Hall, London.

DARROCH, J.N., LAURITZEN, S.L., and SPEED, T.P., 1980: Markov fields and log–linear interaction models for contingency tables. *Annals of Statistics*, **8**, 522–539.

DAVIS, J.A., 1967: A partial coefficient for Goodman and Kruskal's gamma. *Journal of the American Statistical Association*, **62**, 189–193.

DEAN, K., 1989: Conceptual, theoretical and methodological issues in self–care research. *Social Science and Medicine*, **29**, 117–123.

DEAN, K., 1992: Health–related behavior: concepts and methods. In M. Ory, R. Abeles, P. Lipman (eds). *Aging, Health and Behavior.* Sage Publications, Newbury Park, 27–56.

EDWARDS, D., 2000: *Introduction to Graphical Modelling.* 2nd Ed. Springer Pub. Co., New York.

EDWARDS, D. and HAVRÁNEK, T., 1985: A fast procedure for model search in multidimensional contingency tables. *Biometrika*, **72**, 339–351.

KREINER, S., 1987: Analysis of multivariate contingency tables by exact conditional tests: techniques and strategies. *Scandinavian Journal of Statistics*, **14**, 97–112.

KREINER, S., 1992: *Notes on DIGRAM, Version 2.10.* The Danish Institute of Educational Research, Copenhagen, Denmark.

LAURITZEN, S.L., 1996: *Graphical Models.* Oxford University Press, Oxford.

LAURITZEN, S.L. and RICHARDSON, T.S., 2002: Chain graph models and their causal interpretation. *Journal of the Royal Statistical Society, Ser. B,* **64**, 1–28

LAURITZEN, S.L. and WERMUTH, N., 1989: Graphical models for associations between variables, some of which are qualitative and some quantitative. *Annals of Statistics*, **17**, 31–57.

ORY, M., ABELES, R. and LIPMAN, P., (eds) 1992: *Aging, Health and Behavior.* Sage Publications, Newbury Park.

ORY, M. and DeFRIESE, G., (eds) 1998: *Self–Care in Later Life: Research, Program and Policy Perspectives.* Springer Pub. Co., New York.

SMITH, P. and TORREY, B., 1996: The future of the behavioral and social sciences. *Policy Forum, Science,* **271**, 611–612.

SPIRTES, P., GLYMOUR, C., and SCHEINES, R., 1993: *Causation, Prediction, and Search.* Springer Pub. Co., New York.

SUSSER, M. and SUSSER, E., 1996a: Choosing a future for epidemiology: I. Eras and paradigms. *American Journal of Public Health,* **86**, 668–673.

SUSSER, M. AND SUSSER, E. 1996b: Choosing a future for epidemiology: II. From Black Box to Chinese Boxes and eco–epidemiology. *American Journal of Public Health,* **86**, 674–677.

WERMUTH, N. and LAURITZEN, S.L., 1990: On substantive research hypotheses, conditional independence graphs and graphical chain models (with discussion). *Journal of the Royal Statistical Society, Ser. B,* **52**, 21–72.

WHITTAKER, J., 1990: *Graphical Models in Applied Multivariate Statistics.* Wiley, Chichester.

WISTER, A., 1995: The relationship between self–help group participation and other health behaviours among older adults. *Canadian Journal of Community Mental Health,* **14**, 23–37.

Vanessa Didelez
Department of Statistical Science
University College London
Gower Street
London WC1E 6BT
United Kingdom
didelez@stats.ucl.ac.uk

Iris Pigeot

Institute of Statistics	*Bremen Institute for Prevention Research*
University of Bremen	*and Social Medicine*
Postfach 330440	*Linzer Str. 8–10*
28334 Bremen	*28359 Bremen*
Germany	*Germany*

Kathryn Dean and Andrew Wister
Gerontology Research Centre
Simon Fraser University
2800-515 West Hastings Street
Vancouver, British Columbia
Canada, V6B 5K3