International Psychogeriatrics: page 1 of 12 © International Psychogeriatric Association 2011
doi:10.1017/S1041610211000871

Delirium superimposed on dementia: defining disease states and course from longitudinal measurements of a multivariate index using latent class analysis and hidden Markov chains

Antonio Ciampi,[1] Alina Dyachenko,[1] Martin Cole[2] and Jane McCusker[3]

[1] *Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada*
[2] *Department of Psychiatry, St. Mary's Hospital Center and McGill University, Montreal, Canada*
[3] *Department of Clinical Epidemiology and Community Studies, St. Mary's Hospital Center and McGill University, Montreal, Canada*

ABSTRACT

Background: The study of mental disorders in the elderly presents substantial challenges due to population heterogeneity, coexistence of different mental disorders, and diagnostic uncertainty. While reliable tools have been developed to collect relevant data, new approaches to study design and analysis are needed. We focus on a new analytic approach.

Methods: Our framework is based on latent class analysis and hidden Markov chains. From repeated measurements of a multivariate disease index, we extract the notion of underlying *state* of a patient at a time point. The *course* of the disorder is then a sequence of transitions among states. States and transitions are *not* observable; however, the probability of being in a state at a time point, and the transition probabilities from one state to another over time can be estimated.

Results: Data from 444 patients with and without diagnosis of delirium and dementia were available from a previous study. The Delirium Index was measured at diagnosis, and at 2 and 6 months from diagnosis. Four latent classes were identified: *fairly healthy*, *moderately ill*, *clearly sick*, and *very sick*. Dementia and delirium could not be separated on the basis of these data alone. Indeed, as the probability of delirium increased, so did the probability of decline of mental functions. Eight most probable courses were identified, including good and poor stable courses, and courses exhibiting various patterns of improvement.

Conclusion: Latent class analysis and hidden Markov chains offer a promising tool for studying mental disorders in the elderly. Its use may show its full potential as new data become available.

Key words: latent classes, hidden Markov chains, disease course, AIC, BIC, model selection, multiple imputation

Introduction

The study of mental disorders in elderly populations presents substantial challenges due to population heterogeneity, frequent coexistence of more than one disorder in the same patient, and lack of mutually exclusive definitions of the disorders of interest. At a particular point in time, the patient presents a constellation of symptoms and signs, often conceptualized as the manifestation of an underlying *state* of a disorder (e.g. delirium). The evolution of the patient in time is usually referred to as *course*. State and course of a disorder are basic clinical concepts, useful in diagnosis, prognosis and treatment. Data collected on patients with a particular disorder can be used to understand and define these concepts as precisely as possible.

A state may also represent the simultaneous occurrence of more than one disorder (e.g. patients with concomitant delirium and dementia). Therefore, one way of conceptualizing the evolution of such patients is to say that a patient moves in time through the states of a disorder complex. This description is particularly useful if one can identify a finite and preferably small number of highly probable trajectories, each representing a typical *course* of the disorder complex.

In this paper, we consider the situation in which the available data are in the form of repeated measurements of a multivariate index taken on patients at a number of points in time. The values of the multivariate index represent constellations of

Correspondence should be addressed to: Dr. Antonio Ciampi, Department of Epidemiology, Biostatistics and Occupational Health, McGill University, 1020, Pine Ave. West, Montreal, QC, H3A 1A2, Canada. Phone: +1 (514) 398-1584; Fax: +1 (514) 398-4503. Email: antonio.ciampi@mcgill.ca. Received 16 Dec 2010; revision requested 21 Jan 2011; revised version received 8 Apr 2011; accepted 18 Apr 2011.

symptoms and signs. We propose to show that the statistical methodology known as *latent class analysis* can be a powerful tool for abstracting from data meaningful definitions of the state and course of a disorder over time. The use of latent classes enables us to represent a state as a latent class; then, the complex time evolution of the observed symptoms and signs (multivariate index) can be represented as a path through a finite number of latent classes. The key idea is that states are *not* observable; however, we can attach to each patient at each point in time a *probability* of being in a particular state. Similarly, we do *not* observe the transition from one state to another, but we can, through latent class analysis, assign probabilities to each possible transition, and, hence, to each possible trajectory.

Although latent class analysis has a long history (Lazarsfeld and Henry, 1968; Goodman, 1974; McCutcheon, 1987), its use in old age psychiatry is only beginning to be explored, and then only when studying the evolution of a univariate index. Recent examples are the studies by Wilkosza *et al.* (2010) and Terrera *et al.* (2010), both of which focus on describing cognitive decline in an elderly population as measured by a univariate index. The former uses latent class analysis (LCA), the latter relies on growth mixture models (GMM), a related technique. In both LCA and GGM "a given subject may follow a weighted mixture of several entirely different trajectories. The weights represent a set of probabilities, one for each trajectory" (Wilkosza *et al.*, 2010, p. 282). The focus of this paper is on the use of latent class analyses of multivariate indices to describe the evolution of disorders.

To illustrate the proposed approach, we will concentrate on a concrete and challenging example: *delirium* possibly superimposed on *dementia* in the same patient. As described in the DSM-IV, delirium is characterized by acute onset, fluctuating course and potentially reversible disturbances in consciousness, orientation, memory, thought, perception and behavior. Dementia, on the other hand, is characterized by insidious onset, normal level of consciousness and a slowly progressive irreversible decline in mental function. These disabling mental syndromes commonly coexist among older people. Between 25% and 75% of patients with delirium have dementia (Fick *et al.*, 2002); the presence of dementia increases the risk of delirium fivefold (Elie *et al.*, 1998). Moreover, both are probably heterogeneous conditions, each including several different disease states.

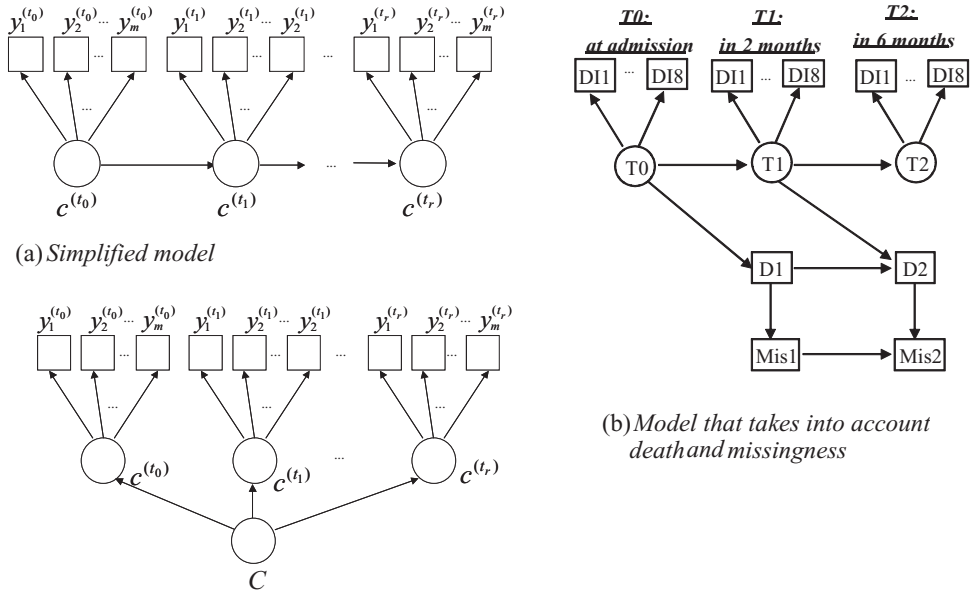
The task of understanding these two conditions, their underlying disease states and interrelationships is formidable. Further studies of phenomenology, epidemiology and prognosis using current methodologies are likely to be of limited value. New methodologies are needed. These new methods may include new measures, new models of pathogenesis (Inouye and Ferrucci, 2006) or novel statistical approaches such as the one described in this paper.

A well-validated instrument used in delirium research is the Delirium Index (DI) (McCusker *et al.*, 2004). The DI was developed to measure the severity of symptoms of delirium. It consists of eight ordinal variables (subscales), each with four symptom levels: absence, low, medium and high. The subscales are defined as: Inattention, Disorganized thinking, Altered level of consciousness, Disorientation, Memory Impairment, Perceptual disturbances, Hyperactivity, Hypoactivity. Clearly, some of the symptoms assessed in the DI are common to both delirium and dementia.

The DI has been used both as a multivariate index and as a univariate score obtained by simply summing scores of the subscales. In a clinical study on delirium in which the multivariate DI was used as the main instrument, Cole *et al.* (2002) attempted to discover different forms of delirium (states of the disorder). Using direct data exploration and cluster analysis, they identified two forms of delirium: hypo-alert and hyper-alert.

The problem of defining delirium course was investigated informally in another study (McCusker *et al.*, 2004) based on the multivariate DI. The evolution in time of the DI subscales over a 12-month period was presented, and it was found that some of the symptoms were more persistent than others. Course, however, was not defined formally. A formal definition of course was developed in a third study, based on the univariate DI score (Sylvestre *et al.*, 2006). A patient's course was represented by the time curve or trajectory of the DI score; finding typical courses was formulated as a problem of clustering trajectories. Using a combination of feature extraction, principal component analysis and k-means clustering, these authors identified five distinct typical courses: Steady, Fluctuating, Worsening, Slow improvement and Fast improvement. The authors did not attempt to work with multivariate DI curves.

The idea of latent classes was absent from the above studies. Instead, these studies detect ("hard") clusters, i.e. groups in data. In general, as a result of the application of cluster analysis to a particular dataset, a patient of the dataset belongs to one and only one of the defined clusters – states or courses in our case. If a new patient has to be assigned to one of the clusters in the above analyses, the clustering framework would only permit assignment of this new patient to one and only one of the clusters. In contrast, latent class analysis aims to discover *non-observable* classes, to which a patient can be assigned



(a) *Simplified model*

(b) *Model that takes into account death and missingness*

(c) *Latent trajectory model*

Figure 1. Graphical representation of latent class models.

with probabilities estimated from the data. The idea of non-observable classes is far better suited than that of clusters to represent the inherent uncertainty in diagnosis of mental disorders.

In this paper we will apply the proposed latent class methodology to a real dataset from a study of delirium conducted at St. Mary's Hospital, a McGill University-affiliated hospital in Montreal, Canada (McCusker *et al.* 2003). First, we describe the latent class model used for the analysis of this dataset, emphasizing the role of the key parameters on which our approach is based. Next, we discuss the estimation of the parameters, outlining the gist of the methodology while avoiding technicalities. Finally, we present the results and their interpretation.

Methods

The proposed model and its parameters

To describe disease states and disease course we propose a statistical model as described by the graphs of Figure 1. We have adhered to the standard convention of graphical models: two or more variables that are *not* directly joined by an arrow are assumed to be conditionally independent given the variables that are joined to them by an arrow (Agresti, 2002). Also, we followed the convention of representing latent variables by circles and manifest variables by squares. Thus, in the case of delirium, the y's denote the components of the DI and the C's denote the latent class variables representing the states of delirium. The sequence of groups of squares denotes repeated measurements of the multivariate DI, and the sequence of circles below represents the sequence of disease states at the time points of the observations. Notice that without the concept of latent class, the only way to represent the dynamics of the illness would be to draw a great deal of arrows joining all y variables, both at each time point and across time points. In contrast, the model proposed here seeks to simplify the representation of the dynamics, by assuming a relatively simple dynamic at a non-observable level; this level provides a simple explanation of the complex evolution of the observable multivariate DI.

A first property of the model apparent from Figure 1a is that we can reconstruct the joint probability distribution of all variables – manifest and latent ones – in terms of a reduced number of probabilistic quantities. Indeed, in order to define the model, we need to specify, at each observation time: (a) the probability distribution of each manifest variable given the corresponding latent class, i.e. given the value of the latent class variable C at the same observation time; (b) the (marginal) probability distributions of the latent class variable; and (c) the transition probability from each latent class at the observation time, to each latent class at the next observation time.

A second property that can be read directly from the graph is the conditional independence, at each observation time, of the DI components, given the latent class. This is the *latent class property* in its classic form (Vermunt, 1997).

A third property represented in the graph is the "Markov property" for the dynamics of the latent

class variables: the probability of being in a disease state at an observation time depends only on the disease state at the previous observation time. In view of the Markov property, the model in Figure 1a is known as a Hidden Markov Chain model (Rabiner, 1989).

Our model therefore rests on a number of assumptions, some more natural than others. We have made them in order to reduce the number of parameters, without sacrificing the essential features we wish to describe. The graph itself suggests how to enlarge the model step by step in case we have enough data to estimate additional parameters. For example, one can add arrows linking the manifest variables, which correspond to the *local dependence* assumption in latent class analysis (Hagenaars, 1990); similarly, if we want to allow a longer "memory" in the dynamics, arrows can be added joining a latent class variable at time t_{i-2} to the latent class variable at t_i.

Two further assumptions, not appearing in the graphs of Figure 1, are *homogeneity* and *stationarity*. By homogeneity we mean that the relationship between manifest and latent variables does not change in time; in other words, the conditional distribution of the manifest variables given the latent class variable is the same at all observation times. The homogeneity assumption plays a key role in developing the description of disease states and course. The stationarity assumption states that the probabilities of the transitions from one state to another during an observation time interval are constant in time. This assumption facilitates the study of disease course: it is also important and easy to test. It should also be noted that the time interval between T_1 and T_2 is the double of the time interval between T_0 and T_1; this is conveniently addressed in our framework by taking the transition matrix between T_1 and T_2 as the square of the transition matrix between T_0 and T_1.

In view of the size of the typical clinical dataset, it may be useful to reduce further the number of model parameters. Without going into too much detail, this can be achieved by introducing some additional "technical' assumptions, which can be translated into logistic regressions to simplify the relationships between manifest and latent variables.

Our model can be improved to include other important aspects of clinical reality. In the case of delirium, two aspects must be considered: (1) a number of patients die in the course of the study; and (2) some patients may not be available for an interview at one or more time points for a variety of uncontrollable reasons. We may decide to restrict the analysis to patients who do not die during the study and are assessed at all time points. However, if we do this, we are left with a much reduced dataset; moreover, if the excluded patients are different from the rest, results may be biased. In Figure 1b we present a model that includes death and missingness with minimal assumptions. The lower part of the graph consists of manifest variables (squares), with D and Mis denoting, respectively, the death and missingness indicators. The additional arrows joining the C's to the D's mean that the probability of dying between two assessments may be a function of the latent class of the patient at the time of the last assessment. There are no arrows joining the latent class variables to the Mis variables; this implies an assumption of random missingness.

To summarize, the parameters of the model in Figure 1a are: (a) the conditional probabilities of the observed values of the multivariate DI components given the latent classes (same at all times under homogeneity); (b) the probabilities of dying given the latent classes during the time interval between two consecutive observation times; (c) the probabilities of loss to follow up given the latent classes during the time interval between two consecutive observation times; and (d) the transition probabilities from latent classes to latent classes for two consecutive observation times (same at all times under stationarity). From these parameters, several quantities of clinical interest can be easily calculated using elementary results of probability theory such as Bayes theorem. As we shall see in the results section, this will permit us to attach probabilities to disease states and disease courses from clinical observations.

For completeness, we note that there is an alternative way of modeling the data. This is represented by the graph of Figure 1c, which represents a model that might be termed a "latent trajectory model". The graph has two layers of latent classes, with the lower level consisting of one latent variable: its latent classes can be directly interpretable as distinct "courses" of the disorder. Appealing as this model may seem, it does miss an important clinical aspect, namely the dependence of the state at a given time on the state at the previous time. It is easy, however, to let the data decide which point of view is preferable, and we have done this, as briefly mentioned in the results section.

Maximum likelihood inference

In order to analyze actual data and interpret the results of the analysis, we need to estimate the parameters of our general model and test various hypotheses about them. To do this, we have relied on maximum likelihood inference, in view of its many desirable properties. This approach can be conveniently applied to our model and data. In

fact, a simple and powerful approach to maximum likelihood inference based on the EM algorithm has been developed (Dempster *et al.*, 1977) for a very broad class of latent class models, including ours. This approach has been implemented in both commercial and free software packages. In this work we have relied on the package LEM (Vermunt, 1997), which is free and in the public domain. In what follows, we briefly describe methods and techniques that are included in LEM.

The estimation of a latent class model is conditional on the number of classes. Usually, this is unknown. We have therefore to consider the number of classes as an index of a super-model family within which we have to select a particular one. In general, an analyst is often interested in comparing a large number of models in order to choose a particular model for in-depth investigation and for drawing substantive conclusions. So, we can distinguish a model selection mode and a model assessment mode of analysis.

When working in model selection mode, the analyst usually relies on a simple approach; using a model performance index that can be computed relatively fast from the data, the best performing model is chosen. The most commonly used performance indices are the Akaike Information Criterion (AIC) (Akaike, 1973) and the Bayesian Information Criterion (BIC) (Schwarz, 1978; Kass and Raftery, 1995). Both of them are likelihood based and are obtained by summing a measure of lack-of-fit and a measure of model complexity. In both cases, model fit is measured by the deviance (-2log-likelihood evaluated at the maximum), while model complexity is measured by a quantity proportional to the number of free parameters; the proportionality coefficient is 2 for the AIC, and the logarithm of the sample size for the BIC. Thus, between two models with the same deviance, these criteria select the one with smaller complexity, in keeping with the general principle of parsimony (Ockham's razor) (Roger, 1976; Vermunt, 1997).

In a model assessment mode, the analyst examines the fit and the properties of the chosen model, often comparing it to other slightly more complex ones, obtained by adding appropriate arrows to the graph of the chosen model. When comparing nested models, likelihood ratio tests are available and can be easily performed. In principle, all assumptions discussed above may be tested using likelihood ratio statistics. In practice, however, this is rarely possible in the latent class context: in fact, in view of the large number of parameters and the sparseness of the data, the χ^2 approximation to the null distribution of the test statistics is of dubious validity. Although one can develop tests based on the bootstrap, we have preferred here to stress the exploratory nature of our modeling effort and to model comparisons based on the BIC. Thus we consider that a model is acceptable if the BICs of models obtained by relaxing simplifying assumptions are larger than the BIC of the model under assessment (Vermunt, 1997).

A simple method for treating missing data is to modify the likelihood by integrating over the missing variables; for example, if some or all of the DI measurements at a particular point in time are missing for a patient, the term in the likelihood corresponding to that patient is modified by summing over all the possible levels of the missing measurements. It should be noted that this is by no means the ideal way to treat missing data, as the underlying assumption is that the patients with missing measurements have exactly the same distribution as those with complete measurements, which implies that data are missing completely at random. Unfortunately, however, alternative approaches are not easily available in popular packages such as LEM.

Results

The dataset and some preliminary decisions

The dataset, well described elsewhere (McCusker *et al.* 2003), was based on an original study of 444 patients conducted at St. Mary's Hospital, a 400-bed primary acute care university-affiliated hospital in Montreal, Canada. Briefly, a study nurse was responsible for patient screening and enrolment in the two studies. Only patients aged 65 years and over who were admitted from the emergency department to medical services were included. Excluded were patients with a primary diagnosis of stroke, those admitted to the oncology unit, those admitted to the intensive care unit or cardiac monitoring unit unless they were transferred to a medical ward within 48 hours of admission, and those who did not speak English or French. Patients were screened for delirium at admission using the Confusion Assessment Method. Patients without delirium at admission were re-screened daily for the following week. Non-delirious subjects were selected from patients screened for delirium but free of this condition. At regular time intervals, a Research Assistant (RA) assessed the patient. The main assessment, from the point of view of this work, was the measurement of the Delirium Index (DI). At enrollment, the RA also collected demographic and clinical data. In particular, the RA measured dementia using the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE), which was dichotomized with a cutoff at 3.5 (if the score was greater than 3.5 the patient was considered

to have dementia). It should be noted that this definition is not longitudinal, but refers to the measurement of the IQCODE at enrollment.

We have worked only with DI measurements taken at enrollment (Time T_0) and at two and six months from enrollment (Time T_1 and T_2 respectively). Table 1 provides the descriptive statistics of the DI measurements and of some demographic and clinical covariates for the 413 patients who had at least two assessments of the DI. Notice that we have coded the DI levels as 0, 1, 2, 3, (from "absence of symptom" to "high level of symptom"), thus implicitly treating the DIs as ordered variables, which we summarize by giving only the means and SD.

Only 188 of 413 patients have complete data at all time points; on the other hand, only 124 patients have missing information beyond baseline. For details, see Table S1 (published as online supplementary material attached to the electronic version of this paper at http://www.journals.cambridge.org/ipg).

Model selection with missing and death indicators

We applied the latent class/HMC modeling techniques described above to the subset of 413 patients who were assessed at times T_0, T_1, and T_2. We included missing and death indicators at the time T_1 and T_2, considering also the possibility that they may be related to other manifest or latent class variables, as in the model shown in Figure 1b. Specifically we assumed: (i) stationarity of the HMC; (ii) homogeneity of the relationship between manifest and latent variables across times T_0, T_1 and T_2; (iii) linearity in the latent variables; and, moreover, (iv) we took the transition matrix between 2 and 6 months to be equal to the square of the transition matrix between enrollment and 2 months.

Table S2 (published as online supplementary material attached to the electronic version of this paper at http://www.journals.cambridge.org/ipg) contains the essential results of our exploration. Our modeling strategy consisted of four steps. At the first step we worked with the model of Figure 1b, varying the number of latent classes. Table S2 shows BIC values for 2, 3, 4 and 5 latent classes: we concluded that the model with 4 latent classes seems to be the best, since it has the smallest BIC.

At the second step we tested a simpler 4 class model, in which the missingness and death indicators are assumed mutually independent and independent of all other variables in the model. At the third step we tested our assumption of stationarity, homogeneity and linearity of the 4 class model. At the fourth step we examined two more complex 4 class models, one which assumes dependence of the missingness indicator on time, and the other which is represented graphically by the latent trajectory model of Figure 1c. As Table S2 shows, all these additional comparisons favor the model of Figure 1b with 4 latent classes.

Model parameters and interpretation

All parameters and its standard errors of our final selected model with 4 latent classes are shown in the Supplementary Appendix (published as online material attached to the electronic version of this paper at http://www.journals.cambridge.org/ipg). Graphical representation of model parameters and interpretation of the main ones is provided below.

Transition matrix

As can be seen from Table S3 (supplementary online material), the transition probabilities have standard deviations between 0 and 0.07. The latent classes are ordered by disorder severity: the healthier patients belong to the lower latent classes, and the sicker to the higher latent classes. The transition matrix shows that the majority of patients tend to remain in the same class during six months from enrollment. More detailed interpretation of these parameters is provided below where we describe the disease course.

Interpreting latent classes

An interpretation of the latent class model is obtained from the distribution of the multivariate DI given the latent classes. This is given in numerical form in Table S4 (supplementary online material), and is presented in graphical form in Figure 2a and 2b. Since both the DI subscales and the latent classes are *ordered*, patients in latent class 1 have less severe DI symptoms than patients in class 2, etc. Figure 2b exploits the order by showing the average values of the DI subscales by latent classes. It should be noted that the Perceptual Disturbance and Hyperactivity subscales do not vary substantially across latent classes; hence these two variables do not help to discriminate among latent classes. The following descriptions are based on Figure 2a.

Class 1 comprises *fairly healthy* patients (*probably no delirium*), who have high probability ($\geq$0.68) of having no problems in focusing attention, in organizing their thoughts, in orientation and in level of activity; however, 28% of them experience a low level of memory problems, and 29% of them have no memory problems at all.

Class 2 comprises *moderately ill* patients with high probability ($\geq$0.80) of no disorganized thinking and no hyper- and hypoactivity; however, focusing attention is concentrated between no symptoms and

Table 1. Descriptive statistics of DI and demographic and clinical covariates, n = 413

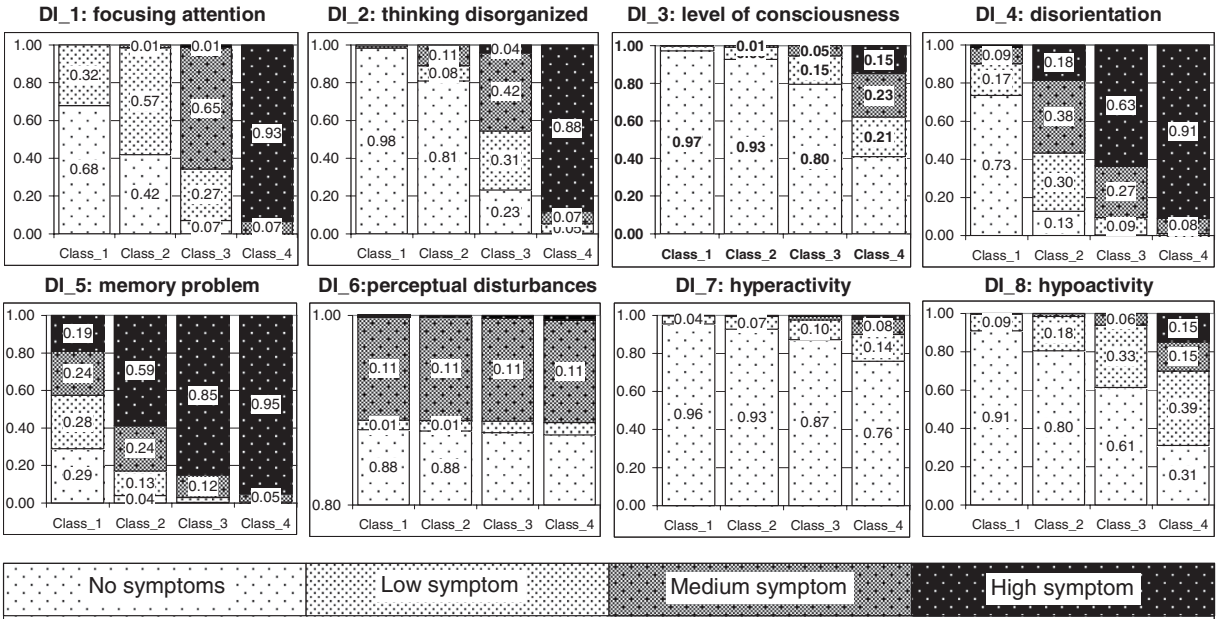
	TIME 0: ADMISSION N = 413		TIME 1: 2 MONTHS N = 249		TIME 2: 6 MONTHS N = 228	
DELIRIUM INDEX VARIABLES	MEAN	STD	MEAN	STD	MEAN	STD
DI_1: focusing attention	1.17	0.93	0.73	0.82	0.74	0.94
DI_2: thinking disorganized	0.82	1.07	0.38	0.8	0.64	1.03
DI_3: altered level of consciousness	0.35	0.69	0.07	0.39	0.04	0.23
DI_4: disorientation	1.7	1.12	1.31	1.21	1.5	1.23
DI_5: memory problem	2.27	0.99	1.96	1.13	2.28	1.03
DI_6: perceptual disturbances	0.31	0.72	0.16	0.53	0.2	0.59
DI_7: hyperactivity	0.2	0.51	0.03	0.19	0.04	0.18
DI_8: hypoactivity	0.47	0.71	0.17	0.47	0.16	0.38

CONTINUOUS COVARIATE	N	MEAN	STD	MIN	MEDIAN	MAX
MMSE	413	16.65	7.35	0	18	29
BARTHEL	413	45.03	29.38	0	44	100
IADL	394	6.94	3.76	0	6	14

COVARIATE	CATEGORY	N	%
Sex	Female	257	62.2
	male	156	37.8
Age	65–74	51	12.3
	74–85	174	42.1
	85+	188	45.5
Living arrangement prior to admission	Home, alone	133	32.2
	Home, with spouse	109	26.4
	Home, with family	56	13.6
	Senior-residence	53	12.8
	Foster-home	20	4.8
	Nursing-home	42	10.2
Dementia	Yes	296	71.7
	No	117	28.3
Severity of illness	Low	60	14.5
	Medium	281	68
	High	72	17.4
Charlson Comorbidity Index	Low	133	32.2
	Medium	166	40.2
	High	114	27.6

DI = Delirium Index; MMSE = Mini-Mental State Examination; Barthel = Barthel Index; IADL = instrumental activities of daily living

(a) *Delirium Index (DI) distribution conditional on four latent classes*



(b) *Average DI distribution conditional on four latent classes*

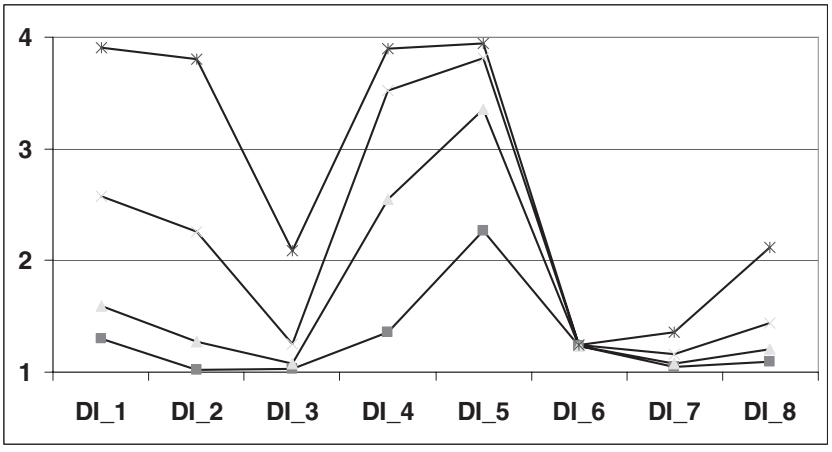


Figure 2. The distribution of the multivariate Delirium Index (DI) given the latent classes: graphical representations.

low symptoms, while disorientation is distributed almost evenly among severity levels, and memory problems have a mostly high severity level (59%) and a medium severity level (24%) (*probably low level of delirium with symptoms of decline of mental function such as found in dementia*).

In Class 3 we find *clearly sick* patients with high level of memory problems (85%), medium and high levels of disorientation (27% and 63% respectively), low and medium level of focusing attention (27% and 65% respectively), low or medium levels of disorganized thinking (31% and 42% respectively), and no or low level of hypoactivity (61% and 33% respectively) (*probably medium level of delirium with medium or high decline of mental function*).

Finally, in Class 4 we find *very sick* patients with high levels of memory problems (95%) and disorientation (91%), high levels of focusing attention (93%), high level of disorganized thinking (88%), and no or low levels of hypoactivity (31% and 39% respectively) (*probably severe form of delirium with high decline of mental function*).

The relationship between delirium latent class and dementia should also be noted. Using the IQCODE to define dementia, our model gives the following percentage of patients with dementia at baseline: 32.4% in Class 1, 59.0% in Class 2, 72.4% in Class 3 and 84.2% in Class 4. Thus, the probability of dementia increases with the order of latent class.

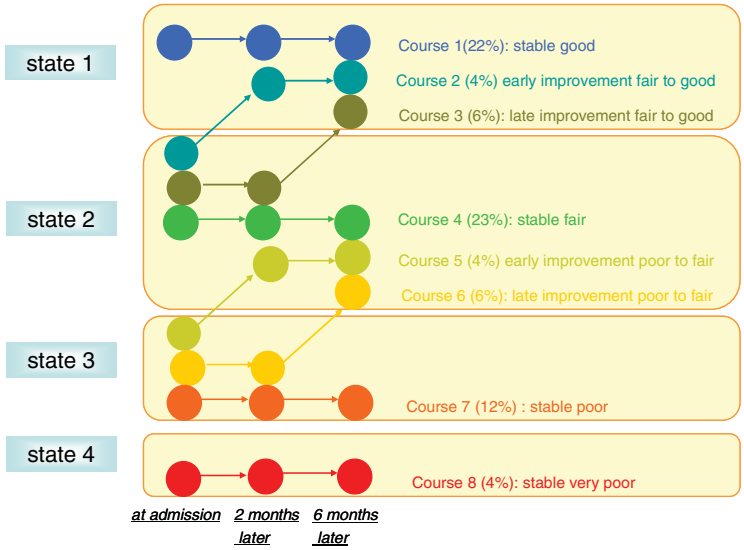


Figure 3. The eight most probable courses of delirium (probability $\geq$ 4%).

Interpreting latent courses

Since the best model is a stationary Hidden Markov Chain (HMC), the initial state probability distribution and the transition matrix entirely define the dynamics of the illness. With the HMC model, every subject has his individual 'course', i.e. his specific path of the HMC. The estimated values of the elements of the HMC transition matrix are shown in the Supplementary Appendix (Table S3). Figure S1 (supplementary online material) represents the same information in graphical form. From a time point to the next, there are $4 \times 4 = 16$ possible transitions (including the transition from a state to itself). The diagonal of Table S3 represents the probabilities of remaining in the same state: these probabilities range from 0.71 to 0.99, indicating a tendency to the status quo. The upper diagonal represents transitions from a better to a worse state; the probabilities range between 0 and 0.03, indicating little tendency to worsening. The lower diagonal represents "improvement", and the probabilities range between 0 and 0.16. On the whole, there are $4 \times 4 \times 4 = 64$ possible distinct paths, i.e. latent courses of the illness; however, most of them are exceedingly rare: for example, 18 latent courses have a probability greater than 0.01, and the probability of any other of the remaining courses is less than 0.03; also, only three latent courses have a probability greater than 0.10, for a cumulative total of 0.58. Figure 3 presents the eight most probable latent courses – those with probabilities $\geq$ 4%.

Posterior classification

At each time point, the model assigns to each of the 413 patients a probability of belonging to each latent class, given the patient's multivariate DI and death indicator. From this it is possible to classify each patient to a unique latent class, the one with the highest posterior probability. Notice that patients with missing information can be classified just as easily as those with complete information. In Table S6 (supplementary online material) we report the number of patients assigned to each class and at each time point.

The 12 panels of Figure 4 show the posterior probabilities of the latent classes for each patient, as they vary over time and across latent classes. In each panel patients are ordered from lowest to highest posterior probabilities of belonging to the latent class to which they have been assigned. The posterior latent class probabilities are color coded: red, yellow, green and blue correspond to the posterior probability of being in latent classes 1, 2, 3 and 4 respectively. It is apparent that the column panels corresponding to latent classes 1 and 4 have one clearly dominant color: this indicates that all or almost all patients belonging to classes 1 and 4 can be classified with a high level of certainty. By contrast, uncertainty is higher for some patients in latent classes 2 and 3.

Model-related frequent questions

In practice, the model can be applied to real data to answer specific questions. A typical clinical question is: "what is the probability distribution of the clinical states of a patient at initial time given the values of the DI scales at initial time?" As an example, consider a patient with the following clinical presentation at initial time:

$y_1^{(t_0)}$ = Medium level of "Focusing attention"
$y_2^{(t_0)}$ = Medium level of "Disorganized thinking"

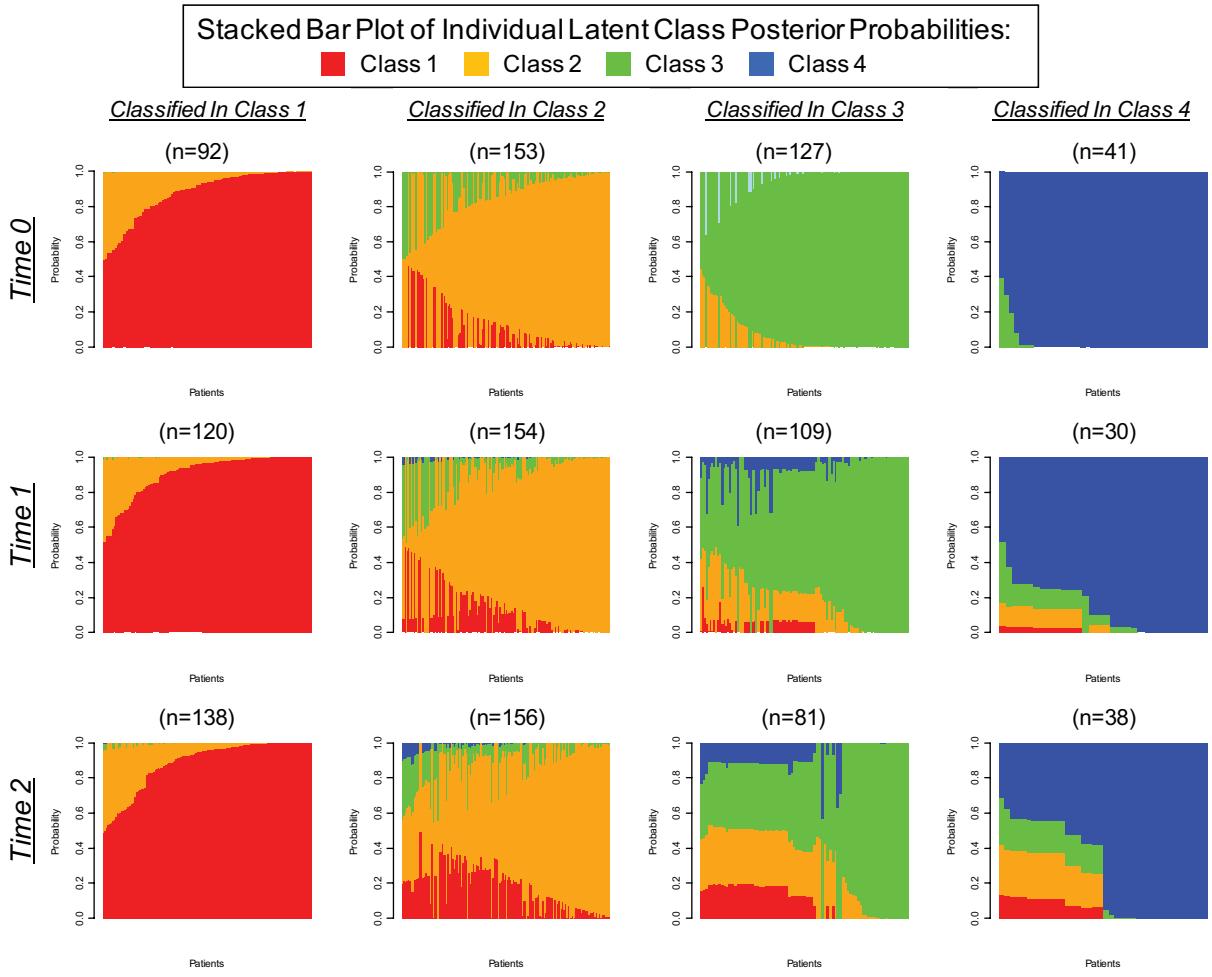


Figure 4. Graphical representation of posterior probabilities of latent class.

$y_3^{(t_0)}$ = Medium level of "Altered level of consciousnesses"
$y_4^{(t_0)}$ = High level of "Disorientation"
$y_5^{(t_0)}$ = High level of "Memory problem"
$y_6^{(t_0)}$ = High level of "Perceptual disturbances"
$y_7^{(t_0)}$ = High level of "Hyperactivity"
$y_8^{(t_0)}$ = High level of "Hypoactivity".

Then from our model we obtain the conditional probabilities of being in states 1, 2, 3, 4, given the above values of the DI scales at t_0, namely 0, 0, 0.75, and 0.25 respectively. Thus the patient is *certainly* in state 3 or 4 and *most probably* in state 3.

At this point one may ask: "what is the probability distribution of the clinical states of a patient in six months if, initially, s/he is in state 3? In state 4?" We find that for patients in state 3 at the initial time, the (conditional) probabilities of being in states 1, 2, 3, and 4 at 6 months are 0.07, 0.44, 0.44, and 0.05 respectively. Similarly for patients in state 4, these probabilities are 0.04, 0.21, 0.42, and 0.33 respectively.

Finally we may ask: "What are the most probable courses of the illness for our patient given the initial clinical presentation?" For a patient with a clinical presentation as above, we find that the three most probable courses are (3,3,3) ($pr = 0.32$), (3,3,2) ($pr = 0.19$), and (3,2,2) ($pr = 0.14$).

Discussion

In this work, we have presented a novel application of latent class analysis to a classical clinical problem: how to define from data the concepts of the state and course of a disease. We have proposed a general framework within the graphical model with latent classes, in which latent classes describe distinct clinical states while the transitions from one clinical state to another over time serves as a basis for defining disease course. We have developed a BIC-based pragmatic strategy for choosing a model within the general framework.

Our framework may be especially useful when studying mental disorders in older people with

conditions that are imperfectly described by a constellation of symptoms and signs. Many mental disorders in old age are of this type. Indeed, mental health clinicians and researchers have attempted to develop disease classifications from data for many years. A well-known result of such a process is the Diagnostic and Statistical Manual of Mental Disorders (DSM). We have attempted to provide a useful tool for work of this type.

Applying this to delirium data, the motivation of this work, has been performed with some success. Indeed, the proposed strategy identified a reasonable model, while providing evidence in support of its fundamental assumptions of homogeneity and stationarity and of the technical assumption of linearity. We found also that our model is preferable to the radically different trajectory model of Figure 1c.

This study has four major strengths. The first is that the four latent classes are interpretable, and the interpretation seems reasonable and consistent with general clinical hypotheses. Class 1 corresponds to the clinical state of a patient who is generally well, except for low levels of impairment as might be expected in an elderly population. Class 2 points to a low level of DI symptoms that are also symptoms of dementia. Classes 3 and 4 seem to describe patients with increasingly severe symptoms of both delirium and dementia. It should be noted that in an earlier work (Cole *et al.*, 2002), simple data exploration led to the identification of two delirium presentations: alert and hypo-alert. Although our classes do not correspond to these two forms, there is evidence that hypo-alert patients are concentrated in Classes 3 and 4, and especially in the latter (three times more hypo-alert patients in Class 4 than in Class 3).

The second strength is that the latent class/hidden Markov chain approach can guide diagnosis and prognosis. As we have shown in a few examples in the Results sectionabove, we can indeed calculate probabilities of initial states and future courses from measurable clinical indices. In our opinion, this is preferable to a categorical approach (e.g. DSM), in which a number of simple rules uniquely assign a patient to one class corresponding to a specific illness.

The third strength of our analysis is the inclusion of patients with incomplete information, which routine techniques would exclude. Finally, the fourth strength of the analysis is the use of powerful graphical representations of the results, which provide substantial clinical insight.

This work has three potential limitations. First, the sample size does not permit us to take full advantage of the model's flexibility. Although our delirium dataset is considered large by clinicians (in fact, these data come from one of the largest datasets available for studying delirium), the models we propose have a fairly large number of parameters. Consequently, it is difficult to separate the uncertainty due to natural variation from that attributable to insufficient sample size. This is apparent, for example, when attempting to classify patients from the posterior probabilities of the latent classes.

The second limitation is the absence of longitudinal data for dementia; indeed, dementia is assessed only at enrolment and not with the same level of depth as delirium. Since delirium and dementia share common symptoms and appear often together in the same patient, this limitation is especially serious. In principle, however, it could be remedied if both illnesses were measured longitudinally and with scales specifically designed for a patient assessment that allows *a priori* that the two illnesses can coexist.

The third limitation is our handling of missing data. We have chosen to use the likelihood integration approach proposed by Vermunt (1997), mainly because it is simple and readily available in LEM, the free software package used for our analysis. We might have used multiple imputation (MI); however, the presence of latent classes and the longitudinal nature of the data makes MI non-trivial both conceptually and computationally.

Our current research aims to overcome some of these limitations. For instance, we are developing a fully Bayesian approach to the estimation of our model parameters which handles missing data while taking into account the presence of latent classes and the longitudinal character of our problem. This proves to be quite complex and computationally heavy, but it will serve as a standard against which to compare some pragmatic, computationally lighter approaches to handling missing data, which we are also investigating.

Another interesting direction of future work is to consider models including more than one latent variable. This would require extending the Hidden Markov Chain treatment of the time evolution to include the description of two concurrent disorders that can be clearly distinguished. The clinical motivation of this effort is clear: for example, we could use two correlated latent variables to describe delirium and dementia simultaneously. However, as mentioned earlier, this would require collecting appropriate data not presently available.

Finally, we plan to extend our approach to allow for the effect of additional covariates on delirium. How does information about the patient, such as gender, age, and general health status, affect the parameters of our model? The graphical representation of our model suggests simple and

intuitive ways to introduce the effect of these variables on the latent classes. Details, however, require further investigation.

We must add that there are other limitations to the present findings that cannot be addressed by improvements of statistical methodology alone. Some clinicians may not find our definitions of classes entirely convincing. For example, to some, the probability of 68% of no impairment associated with Class 1, which we have referred to as "high", may not appear all that high. Clearly, for a classification to be useful as a current research tool, these details should be addressed and definitions should be arrived at by clinical consensus. We therefore wish to emphasize that the progress in this work is above all conceptual. To obtain substantial clinical progress, this work should be followed by studies in which main findings can be replicated or incrementally modified. Further studies, as we have mentioned, should also attempt to follow longitudinally the cognitive decline associated with dementia.

In conclusion, we hope that this work may contribute towards the development of statistical tools that help maximize the ability of geriatric psychiatry researchers to extract information from clinical data, and develop empirical classifications for old age mental disorders and their courses. Tools other than analytical will, of course, need to be developed, such as appropriate study designs for collecting data that appropriately reflect a target population.

Conflict of interest

None.

Description of authors' roles

Antonio Ciampi proposed the idea of using latent class methodology. Alina Dyachenko contributed substantially to the implementation of the methodology and performed the calculations. Martin Cole and Jane McCusker contributed their clinical and methodological expertise at each step of the development, providing constructive criticism, and making the data available.

References

Agresti, A. (2002). *Categorical Data Analysis*. New Jersey: John Wiley & Sons.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory* (pp. 267–281). Budapest: Akademiai Kiado.

Cole, M. G., McCusker, J. and Dendukuri, N. (2002). Symptoms of delirium among patients with or without dementia. *Journal of Neuropsychiatry and Clinical Neurosciences*, 14, 167–175.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39, 1–38.

Elie, L. M., Cole, M. G., Primeau, F. J. and Bellavance, F. (1998). Delirium risk factors in elderly hospitalized patients. *Journal of General Internal Medicine*, 13, 204–212.

Fick, D. M., Agostini, J. V. and Inouye, S. K. (2002). Delirium superimposed on dementia: a systematic review. *Journal of the American Geriatrics Society*, 50, 1723–1732.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61, 215–231.

Hagenaars, J. A. (1990). *Categorical Longitudinal Data*. Newbury Park, CA: Sage.

Inouye, S. K. and Ferrucci, L. (2006). Elucidating the pathophysiologyof delirium and the interrelationship of delirium and dementia. *Journals of Gerontology, Series A: Medical Sciences*, 61A, 1277–1280.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.

Lazarsfeld, P. F. and Henry, N. W. (1968). *Latent Structure Analysis*. Boston: Houghton Mifflin.

McCusker, J., Cole, M. G., Dendukuri, N., Han, L. and Belzile, E. (2003). The course of delirium in older medical inpatients: a prospective study. *Journal of General Internal Medicine*, 18, 696–704.

McCusker, J., Cole, M. G., Dendukuri, N. and Belzile, E. (2004). The Delirium Index – a measure of severity of delirium: new findings on reliability, validity and responsiveness. *Journal of the American Geriatrics Society*, 52, 1744–1749.

McCutcheon, A. C. (1987). *Latent Class Analysis*. Beverly Hills: Sage Publications.

Rabiner, L. R. (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, 77, 257–286.

Roger, A. (1976). *Ockham's Razor: A Historical and Philosophical Analysis of Ockham's Principle of Parsimony*. Champaign-Urbana, IL: University of Illinois.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.

Sylvestre, M. P., McCusker, J., Cole, M., Regeasse, A., Belzile, E. and Abrahamowicz, M. (2006). Classification of patterns of delirium severity scores over time in an elderly population. *International Psychogeriatrics*, 18, 667–680.

Terrera, G. M., Brayne, C., Matthews, F. and the CC75C Study Collaboration Group. (2010). One size fits all? Why we need more sophisticated analytical methods in the explanation of trajectories of cognition in older age and their potential risk factors. *International Psychogeriatrics*, 22, 291–299.

Vermunt, J. K. (1997). *Log-linear Models for Event Histories*. Thousand Oaks, CA: Sage.

Wilkosza, P. A. *et al.* (2010). Trajectories of cognitive decline in Alzheimer's disease. *International Psychogeriatrics*, 22, 281–290.