BULLETIN *of the*
MALAYSIAN MATHEMATICAL
SCIENCES SOCIETY
http://math.usm.my/bulletin

Bull. Malays. Math. Sci. Soc. (2) **36***(3) (2013), 637–655*

A Multistate Transition Model for Analyzing Longitudinal Depression Data

[1]M. ATAHARUL ISLAM, [2]RAFIQUL I. CHOWDHURY AND [3]SHAHARIAR HUDA
[1]Department of Statistics and OR, King Saud University, P. O. Box 2455, Riyadh 11451 Saudi Arabia
[2]Department of Epidemiology and Biostatistics, Western University, London, ON, Canada
[3]Department of Statistics and OR, Kuwait University, Kuwait
[1]mataharul@yahoo.com, [2]chowdhuryri@yahoo.com, [3]shahariarhuda@yahoo.com

Abstract. In longitudinal data analysis, there are many practical situations where we need to deal with transitions to a number of states and which are repeated over time generating a large number of trajectories from beginning to end of the study. This problem becomes increasingly difficult to model if the number of follow-ups is increased for a set of longitudinal data. A covariate-dependent Markov transition model is proposed using the logistic link function for polytomous outcome data. A generalized and more flexible approach of constructing the likelihood function for the first or higher order is demonstrated in this paper to deal with the branching of a number of transition types starting from no depression at the beginning of the study. The proposed method can be employed to resolve a long-standing problem in dealing with modeling for transitions, reverse transitions and repeated transitions by reducing the number of trajectories to a large extent resulting in estimating relatively few parameters. The problem of depression in elderly, in terms of short and long-term health and economic consequence, needs to be assessed more critically. This study uses the longitudinal data from the six waves of the Health and Retirement Survey to examine the transition to depression, reverse transition from depression to no depression and also repeated transition from no depression to depression after experiencing a reverse transition during a study period. The results indicate that age is negatively associated with reverse and repeated transitions, gender is negatively associated with transition and reverse transition indicating that females are more likely to experience both. The proposed method clearly provides a wider range of useful information in revealing the dynamics of the depression pattern among elderly.

2010 Mathematics Subject Classification: 62H20, 62F03, 62F10, 62N03

Keywords and phrases: Markov chain, multistate model, logistic regression, depression, and transitions.

1. Introduction

The disease process of depression involves several stages over a long period of time. To understand this process we need to examine the sequence of events during subsequent follow-ups. The Markov model is a natural choice for the study of depression due to its underlying

Communicated by Anton Abdulbasah Kamil.
Received: January 13, 2012; *Revised:* July 16, 2012.

property of dependence in consecutive follow-ups. The higher order Markov model can be used to investigate whether the relationship between outcomes can be explained in terms of longer term effects of depression due to recurrence of depressive episodes during long spells of observation. It is important to understand how the transitions among the different states of depression take place and how the covariates influence these transitions. Markov chain is one class of statistical models which is used to identify the characteristics of these types of transitions. The Markov chain models for discrete variate time series appear to be restricted due to over-parameterization and several attempts have been made to simplify the application of Markov chain models. There are different approaches of employing Markov chain models. Raftery [43], Raftery and Tavare [42], and Berchtold and Raftery [6] addressed one such area of problems in estimating transition probabilities following the work of Pegram [40]. This area of research, popularly known as the mixture transition distribution (MTD), deals with modeling of high-order Markov chains for a finite state space. In another development, Albert [1] proposed a finite Markov chain model for analyzing sequences of ordinal data from a relapsing remitting of a disease. In addition, Albert and Waclawiw [2] developed a class of quasi-likelihood models for a two state Markov chain with stationary transition probabilities for heterogeneous transitional data. However, these models deal with only estimation of transition probabilities. In another attempt, Buhlmann and Wyner [9] developed a class of stationary variable length Markov chains on a finite space. They demonstrated that these processes essentially follow higher order Markov chains with potential memory of variable length with applications to information theory and machine learning.

Regier [44] proposed a model for estimating odds ratio from a two state transition matrix. A grouped data version of the proportional hazards regression model for estimating computationally feasible estimators of the relative risk function was proposed by Prentice and Gloeckler [41]. The role of previous state as a covariate was examined by Korn and Whittemore [32]. Wu and Ware [49] proposed a model which included accumulation of covariate information as time passes before the event and considered occurrence or non-occurrence of the event under study during each interval of follow up as the dependent variable. The method could be used with any regression function such as the multiple logistic regression model and also the one suggested by Prentice and Gloeckler [41]. Kalbfleisch and Lawless [29] proposed the models for continuous time. They presented procedures for obtaining estimates for transition intensity parameters in homogeneous models. For a first order Markov model, they introduced covariate dependence of log-linear type. Due to complexity in the formulation of the underlying models, none of these models could be generalized to higher order.

More recently, another class of models has emerged for analyzing transition models with serial dependence of the first or higher orders on the basis of the marginal mean regression structure models. Azzalini [4] introduced a stochastic model, more specifically, a first order Markov model, to examine the influence of time-dependent covariates on the marginal distribution of the binary outcome variables in serially correlated binary data. Markov chains are expressed in transitional form rather than marginally and the solutions are obtained such that covariates relate only to the mean value of the process, independent of association parameters. Following Azzalini [4], Heagerty and Zeger [22] presented a class of marginalized transition models (MTM) and Heagerty [21] proposed a class of generalized MTMs to allow serial dependence of first or higher order. These models are computationally tedious and the form of the serial dependence is quite restricted. If the regression parameters are strongly influenced by inaccurate modeling for serial correlation then the MTMs can result in misleading conclusions. Heagerty [21] provided derivatives for score and information computations. Lindsey and Lambert [36] examined some important theoretical aspects concerning the use of marginal models and demonstrated that there are serious limitations such as: (i) produce profile curves that does not represent any possible individual, (ii) may show that a treatment is better on average when, in reality, it is poorer for each individual subject, (iii) generate complex and implausible physiological explanations with underdispersion in subgroups and problems associated with no possible probabilistic data generating mechanism.

In recent years, there has been a great deal of interest in the development of multivariate models based on the Markov Chains. These models have wide range of applications in the fields of reliability, economics, survival analysis, engineering, social sciences, environmental studies, biological sciences, etc. Muenz and Rubinstein [38] employed logistic regression models to analyze the transition probabilities from one state to another but still there is a serious lack of a general methodology for analyzing transition probabilities of higher order Markov models. In a higher order Markov model, we can examine some inevitable characteristics that may be revealed from the analysis of transitions, reverse transitions and repeated transitions. Islam and Chowdhury [27] extended the model for higher order with covariate dependence for binary or polytomous outcomes. It is noteworthy that the covariate dependent higher order Markov models can be used to identify the underlying factors associated with such transitions.

In this study, it is aimed to provide a comprehensive covariate-dependent Markov model for higher order. The proposed model is a further generalization of the models suggested by Muenz and Rubinstein [38] and Islam and Chowdhury [27] in dealing with event history data. Lindsey and Lambert [36] observed that the advantage of longitudinal repeated measures is that one can see how individual responses change over time. They also concluded that this must generally be conditioned on the previous history of a subject, so that it usually would seem that analyses that concentrate on the marginal aspects of models are discarding important information, or not using it efficiently. Lee and Nelder [34] concluded that the conditional models are of fundamental interest and marginal predictions can be made from conditional models. The transition models are developed on the basis of conditional models employing Markovian assumptions and provide more realistic formulation as compared to the marginal models such as the generalized estimating equations. As the classical marginal models (such as GEE) may pose difficulty with the selection of underlying correlation structure [28], which may or may not be functions of the marginal. Another alternative is the subject-specific models taking into consideration the random effects by allowing random effect terms in the linear predictor [7]. Lee and Nelder [34] observed that conditional models are fundamental and the advantages of conditional models over marginal models are obvious because marginal predictions can be made from conditional models. In our study, we have considered polytomous logit link function for m outcome categories (categories are denoted by $u = 0, 1, 2, \ldots, m-1$). We defined the set of repeated outcome variables for subject i at the j-th follow-up as a vector, $(Y_{i1}, Y_{i2}, ..., Y_{ij})$, where $i = 1, .2, \ldots, n$ and $j = 1, 2, \ldots, J_i$. Here Y_{ij} corresponds to the response at time t_{ij}. Then the random effect model can be denoted by [32]:

$$\ln\left(\frac{\pi_{iju}}{\pi_{ij0}}\right) = \theta_u + x'_{ij}\beta_u + \nu_{iu},\ u = 1, 2, ..., m-1,$$

where $\pi_{iju} = P(Y_{ij} = u)$ are probabilities of outcomes, θ_u and $\beta_u = (\beta_{u1}, ..., \beta_{up})'$ are fixed effects. The random component $v_i \sim N(0, \Sigma)$, where Σ is unstructured covariance matrix. Again the assumption on covariance structure remains a problem in this type of mixed effect models.

It has been observed that the models based on generalized estimating equations provide very attractive and useful results but the estimates are often inefficient [16, 10]. In contrary to the classical GEE or the random effect model described above, the proposed transitional models introduced in Section 3, provides a more general and flexible set up for addressing the issue of analyzing repeated polytomous outcomes emerging from longitudinal studies. We do not need to impose any correlation or covariance structure and unlike the other alternatives the transition models can be employed to examine the relationship between previous and current outcomes in terms of risk factors or covariates of interest.

Our proposed model is based on conditional approach and uses the event history efficiently. Furthermore, using the Chapman-Kolmogorov equations, the proposed model introduces an improvement over the previous methods in handling runs of events which is common in longitudinal data. It is noteworthy that this paper shows that we can express the conditional probabilities in terms of the transition probabilities generated from Markovian assumptions. A general procedure is developed comprehensively in this paper to propose the estimation procedure for Markov models of any order. The proposed model and inference procedures are simple and the covariate dependence of the transition probabilities of any order can be examined without making the underlying model complex. Another advantage of the model lies in the fact that the estimation and test procedures for both the specific parameter of interest and the overall model remain easy for practical applications for any longitudinal data.

In a recent study, Harezlak *et al.* [20] proposed a Markov model consisting of three states, non-diseased, diseased and dead, and estimated the transition hazard parameters using the maximum likelihood approach for analyzing dementia data. Their model does not include any reverse transition or repeated transition. However, instead of considering death as missing, they considered it as a separate state and thus the results became more meaningful. Islam and Singh [25], Islam [26], and Islam *et al.* [24] demonstrated the analysis of transitions, reverse transitions and repeated transitions for longitudinal data. They employed hazards model for continuous time data with discrete state outcomes. Islam *et al.* [24] indicated the possibility of using logistic link functions for similar analysis. In the previous works, the transitions were considered to follow hidden Markov properties of first or higher orders. In other words, we needed to strictly observe consecutive follow-ups to obtain the history of events. In that case, the branching of possible occurrence of the events would produce a large number of models based on the past history. This paper introduces a simplified approach based on the Chapman-Kolmogorov equations, to formulate the models for transitions, reverse transitions and repeated transitions. The likelihood function proposed in this paper makes the longitudinal problems for transitions among a number of states much more flexible and thus provides a more useful generalization of the estimation procedure for repeated measures. In addition, the confirmed deaths during consecutive follow-ups are considered from both non depressed and depressed states so that the deaths are not included as censored because there might be both qualitative and quantitative differences between other types of censoring and death, particularly in analyzing depression.

The emerging issues concerning the analysis of depression among elderly include the identification of risk factors from longitudinal studies. The Health and Retirement Study (HRS) conducted in the USA provides an opportunity to examine the follow-up data on elderly population covering six consecutive follow-ups over a period of 10 years, each 2 years apart. In this paper, we have proposed a covariate dependent Markov model for three states, no depression, depression and death. We have considered death as a separate state in order to provide a more useful procedure for analyzing the transition from no depression to depression as well as for analyzing reverse transition from depression to no depression and repeated transition for those who made a reverse transition at a previous stage.

2. The covariate dependent Markov model for transition, reverse transition and repeated transition

Let us consider the set of repeated outcome variables for subject i at the j-th follow-up as a vector, $(Y_{i1}, Y_{i2}, ..., Y_{ij})$, where i = 1,.2,...,n and j =1,2,..., J_i. Here Y_{ij} corresponds to the response at time t_{ij}. Denote by $H_{ij} = \left\{Y_{ij'},\, j' = 1,2,...,j-1\right\}$ the past history of subject i at the j-th follow-up. Let us denote $Y_{ij_1-1} = s_{j_1-1}$, $s_{j_1-1} = 0,1,2,...,m-1$; $Y_{ij_1} = s_{j_1}$, $s_{j_1} = 0,1,2,...,m-1$ and so on for occurrence of events for the i-th subject at times t_{j_1-1} and t_{j_1} respectively where Y_{ij_1-1} =0 and Y_{ij_1} =0 indicate that no events occurred. The first order Markov model for an event at time t_{j_1} conditional upon an event at time t_{j_1-1} is

$$P(Y_{ij_1} \left| Y_{ij_1-1}\right.) \,. \tag{2.1}$$

We consider here m possible outcomes for the response variable at each follow-up. The possible outcome categories are $0,1,...,m-1$ for the response variable, Y. The conditional probability of $Y_{ij_1} = s_{j_1}$ $(s_{j_1} = 0,1,2,...,m-1)$ at time t_{j_1} given $Y_{ij_1-1} = s_{j_1-1}$, $(s_{j_1-1} = 0,1,2,...,m-1)$ at time t_{j_1-1} is $\pi_{s_{j_1}|s_{j_1-1}} = P(Y_{ij_1} = s_{j_1} \left| Y_{ij_1-1} = s_{j_1-1}\right.)$ for any s_{j_1-1}, where

$$\sum_{s_{j_1}=0}^{m-1} \pi_{s_{j_1}|s_{j_1-1}} = 1, \quad s_{j_1-1} = 0,1,2,...,m-1. \tag{2.2}$$

We can show that

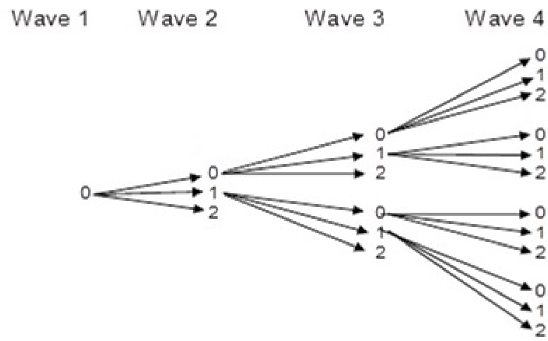
$$\begin{aligned} \pi_{s_{j_1-1}s_{j_1}} &= P(Y_{ij_1} = s_{j_1}, Y_{ij_1-1} = s_{j_1-1}) \\ &= P(Y_{ij_1} = s_{j_1} \left| Y_{ij_1-1} = s_{j_1-1}\right.) \times P(Y_{ij_1-1} = s_{j_1-1}) \\ &= \pi_{s_{j_1}|s_{j_1-1}} . \pi_{s_{j_1-1}} \end{aligned} \tag{2.3}$$

where $\pi_{s_{j_1-1}} = P(Y_{j_1-1} = s_{j_1-1})$.

Hence, we can express the first order transitions during the following time intervals: $(t_{j_1-1}, t_{j_1}), (t_{j_2-1}, t_{j_{12}}), ..., (t_{j_k-1}, t_{j_k})$. The transition probabilities can be shown as:

$$\begin{aligned} \pi_{s_{j_u-1}s_{j_u}} &= P(Y_{ij_u} = s_{j_u}, Y_{ij_u-1} = s_{j_u-1}) \\ &= P(Y_{ij_u} = s_{j_u} \left| Y_{ij_u-1} = s_{j_u-1}\right.) \times P(Y_{ij_u-1} = s_{j_u-1}) \\ &= \pi_{s_{j_u}|s_{j_u-1}} . \pi_{s_{j_u-1}}, \qquad u = 2,...,k. \end{aligned} \tag{2.4}$$

Now let us consider $m = 3$. The states are: no depression (0), depression (1) and death (2). Figure 1 displays the flow of transitions, reverse transitions and repeated transitions, starting at the baseline from the state of no depression.



0 = No Depression, 1= Depression, and 2 =Death

Figure 1. Flow chart for different transition types

We can define the following probabilities for k time points, using the Chapman-Kolmogorov quations [44] and also using Equation (2.1). The joint probability of s_{j_1-1} at time t_{j_1-1} and s_{j_1} at time t_{j_1} is, where t_{j_1-1} is the time of follow-up just prior to t_{j_1}:

$$\pi_{s_{j_1-1}s_{j_1}} = \pi_{s_{j_1}|s_{j_1-1}} . \pi_{s_{j_1-1}}. \tag{2.5}$$

Similarly, we can obtain the joint probability based on probability of transition from s_{j_1-1} at time t_{j_1-1} to s_{j_1} at time t_{j_1} and then from s_{j_1} at time t_{j_2-1} and s_{j_2} at time t_{j_2} as follows, and assuming no other events occurred in between other than remaining in the same state in subsequent follow-ups until the time of making a transition: Here

$$\pi_{s_{j_1}s_{j_2}s_{j_3}} = P(Y_{ij_3} = s_{j_3} \left| Y_{ij_3-1} = s_{j_2} \right.) \times P(Y_{ij_2} = s_{j_2} \left| Y_{ij_2-1} = s_{j_1} \right.) \times P(Y_{ij_1} = s_{j_1}).$$

Here, we assumed that the probability remains same for remaining in the same state after a transition until the time of the subsequent transition. This implies that we expect transitions only at time intervals $(t_{j_1-1}, t_{j_1}), (t_{j_2-1}, t_{j_{12}}), ..., (t_{j_k-1}, t_{j_k})$ and transitions do not occur in between. In other words, we can express the Equations (2.4) equivalently as:

$$\begin{aligned} \pi_{s_{j_1}s_{j_2}s_{j_3}} &= P(Y_{ij_3} = s_{j_3} \left| Y_{ij_3-1} = s_{j_2} \right.) P(Y_{ij_2} = s_{j_2} \left| Y_{ij_2-1} = s_{j_1} \right.) \times P(Y_{ij_2-1} = s_{j_1}) \\ &= \pi_{s_{j_3}|s_{j_2}} \times \pi_{s_{j_2}|s_{j_1}} \times \pi_{s_{j_1}}. \end{aligned} \tag{2.6}$$

For k time points of occurrence of events, this can be generalized as follows:

(2.7)

$$\begin{aligned} \pi_{s_{j_1}...s_{j_k}} &= P(Y_{ij_k} = s_{j_k} \left| Y_{ij_k-1} = s_{j_{k-1}} \right.) \times ... \times P(Y_{ij_2} = s_{j_2} \left| Y_{ij_2-1} = s_{j_1} \right.) \times P(Y_{ij_2-1} = s_{j_1}) \\ &= \pi_{s_{j_k}|s_{j_{k-1}}} \times ... \times \pi_{s_{j_2}|s_{j_1}} \times \pi_{s_{j_1}} \\ &= \pi_{s_{j_1}} \prod_{u=2}^{k} \pi_{s_{j_u}|s_{j_{u-1}}}. \end{aligned}$$

Let us illustrate the above through a simple example with only two states. Then the states $s_{j_1}, s_{j_2}, s_{j_3}$ and s_{j_4} can have only two values, 0 or 1, and a follow-up at time points starting from t_0 to t_6 may result in $2^7 = 128$ possible paths. A typical path may be as 0-0-0-1-1-0-0. It involves a large number of potential transition types as the time points for follow-up increase due to the inclusion of runs of 0 or 1 before making a transition. For instance, in the first transition type, 0 was observed in subsequent times t_0, t_1, t_2 and then 1 was observed at times t_3, t_4 and subsequently 0 again at follow-up times t_5, t_6. In terms of transition, we observe here only two transitions, 0-1 (during times t_2, t_3) and then 1-0 (during times t_4, t_5). These transitions can be denoted as $s_{j_1} \to s_{j_2}$ (0-1), $s_{j_2} \to s_{j_3}$ (1-0). In other words, the first transition for any individual is denoted as $s_{j_1} \to s_{j_2} (u \to v)$ which is called a transition, and the second one is $s_{j_2} \to s_{j_3}$ $(v \to w)$ which is a reverse transition and the third one can be extended as $s_{j_3} \to s_{j_4} (w \to s)$ which is a repeated transition in the above illustration. Figure 2 shows the possible transition types based on the above example.

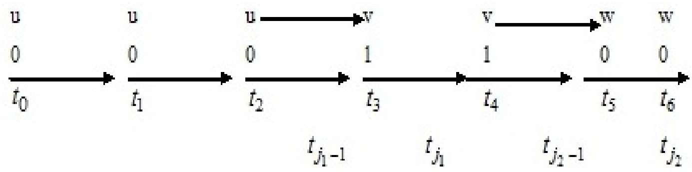


Figure 2. Flow chart for transitions for the typical path

Now let us develop the link function for the conditional probabilities. The covariate vector for individual i is $X_i = [1, \quad X_{i1}, \quad ...., \quad X_{ip}]$ and the corresponding vector of parameters for transition from $s_{j_{u-1}}$ to s_{j_u} is $\beta'_{s_{j_{u-1}} s_{j_u}} = \left[\beta_{s_{j_{u-1}} s_{j_u} 0}, \ \beta_{s_{j_{u-1}} s_{j_u} 1}, \ ...., \ \beta_{s_{j_{u-1}} s_{j_u} p} \right]$,

As there are m possible outcomes for the response variable Y at each follow-up, we can express the transition from u to v in terms of polytomous logistic regression as follows:

$$\pi_{s_{j_u} | s_{j_{u-1}}} (X) = P(Y_{j_u} = s_{j_u} | Y_{j_{u-1}} = s_{j_{u-1}}, X) = \frac{e^{g_{s_{j_{u-1}} s_{j_u}}(X)}}{\sum_{s'_{j_u}=0}^{m-1} e^{g_{s_{j_{u-1}} s'_{j_u}}(X)}} \tag{2.8}$$

where $s_{j_u} = 0, 1, ..., m$,

$$g_{s_{j_{u-1}} s_{j_u}}(X) = \begin{cases} 0, & if \ s_{j_u} = 0 \\ \ln \left[\dfrac{P(Y_{j_u} = s_{j_u} | Y_{j_u - 1} = s_{j_{u-1}}, X)}{P(Y_{j_u} = 0 | Y_{j_{u-1}} = s_{j_{u-1}}, X)} \right] & , \qquad if \ s_{j_{u-1}} = 1, ..., m-1, \end{cases}$$

with $g_{s_{j_{u-1}} s_{j_u}}(X) = \beta_{s_{j_{u-1}} s_{j_u} 0} + \beta_{s_{j_{u-1}} s_{j_u} 1} X_1 + ... + \beta_{s_{j_{u-1}} s_{j_u} p} X_p$.

3. Estimation

The model developed in Section 2 considers m possible outcomes $0, 1, ..., m-1$. Here the transitions start from the state s_{j_1}. To examine the transitions in the subsequent times, we observe the conditional probabilities starting from the initial state and the corresponding probability. Then the transitions to different state categories are observed in subsequent follow-ups. The likelihood function for n individuals with the i-th individual $(i = 1, 2, ..., n)$

having (k-1) possible transitions specified by k time points $t_{j_1}, t_{j_2}, ..., t_{j_k}$ can be expressed as:

$$L = \prod_{i=1}^{n} \left[\pi_{s_{j_1}}(X_i) \prod_{u=2}^{k} \left\{ \pi_{s_{j_u}|s_{j_{u-1}}}(x_i) \right\}^{\delta_{is_{j_{u-1}}s_{j_u}}} \right] \tag{3.1}$$

where $\delta_{is_{j_{u-1}}s_{j_u}} = 1$, if a transition of the type $s_{j_{u-1}} \to s_{j_u}$ occurs to the i-th individual during the time interval $(t_{j_{u-1}}, t_{j_u})$, $\delta_{is_{j_{u-1}}s_{j_u}} = 0$, otherwise.

The log likelihood function is given by

$$\ln L = \sum_{i=1}^{n} \left[\ln \pi_{s_{j_1}}(x_i) + \sum_{u=2}^{k} \delta_{is_{j_{u-1}}s_{j_u}} \ln \left\{ \pi_{s_{j_u}|s_{j_{u-1}}}(x_i) \right\} \right]. \tag{3.2}$$

Differentiating (3.2) with respect to the parameters we obtain the log-likelihood equations and solving the equations, we obtain the maximum likelihood estimates of the parameters for transition type $s_{j_{u-1}} \to s_{j_u}$as follows:

$$\frac{\partial \ln L}{\partial \beta_{s_{j_{u-1}}s_{j_u}q}} = \sum_{i=1}^{n} \sum_{j_1=1}^{j_2} \delta_{is_{j_{u-1}}s_{j_u}} X_{qi} (1 - \pi_{s_{j_u}|s_{j_{u-1}}}(x_i)) = 0 \quad , q = 0, 1, \ldots, p. \tag{3.3}$$

It is noteworthy that the first part of Equation 3.2 does not contain any parameter of interest for the probabilities of transitions we defined in 2.8.

There is no closed form solution of these equations. So we need to employ an iterative process based on the Newton-Raphson method. For a quick convergence, we need to use the second derivatives (which will provide the elements for the observed information matrix after multiplication by -1),

$$\frac{\partial^2 \ln L}{\partial \beta_{s_{j_{u-1}}s_{j_u}q} \partial \beta_{s_{j_{u-1}}s_{j_u}q'}} = - \sum_{i=1}^{n} \delta_{is_{j_{u-1}}s_{j_u}} X_{q'i} X_{qi} \pi_{s_{j_u}|s_{j_{u-1}}}(x_i)(1 - \pi_{s_{j_u}|s_{j_{u-1}}}(x_i)), \tag{3.4}$$

$$\frac{\partial^2 \ln L}{\partial \beta_{s_{j_{u-1}}s_{j_u}q} \partial \beta_{s_{j_{u-1}}s'_{j_u}q'}} = - \sum_{i=1}^{n} \delta_{is_{j_{u-1}}s_{j_u}} X_{q'i} X_{qi} (\pi_{s_{j_u}|s_{j_{u-1}}}(x_i))(\pi_{s'_{j_u}|s_{j_{u-1}}}(x_i))$$

where $u = 2, \ldots, k; q, q' = 0, \ldots, p; s_{j_{u-1}} = 0, 1, ..., m-1; s_{j_u} = 1, ..., m-1$.

The models proposed here are different from the independent logistic regression models for 0-1, 0-2, 0-1-0, 0-1-2, 0-1-0-1, 0-1-0-2 because the proposed models consider the event history while the independent logistic regression models do not take account of progression of events completely. Let us consider the transitions, for instance, 0-1, 0-1-0, 0-1-2, 0-1-0-1 and 0-1-0-2, where all these transitions have 0-1 as the first transition, and if we consider these separately to fit distinct models, then the sample size for the first transition is reduced resulting in biased estimates of the parameters due to ignoring a substantial proportion of first transition of the type 0-1 which progressed further with other transitions, reverse transitions and repeated transitions.

We can also define the likelihood ratio statistic for the overall model based on Equations 3.2–3.4 for testing the hypotheses concerning the parameters of the models for different types of transitions. We can employ the BIC for identifying the best model out of the potential alternatives using (3.2): $BIC = -2\ln L_r + r\ln n$, where $L_r$is the maximized likelihood, r =number of parameters in the model, and n =sample size.

4. Application to depression data

During the past decades, many developed and developing countries experienced steady increase in their elderly population. With an increase in proportion of elderly population, the focus of public health also needs major adjustments in health policies in order to face challenges due to change in age composition. One such major problem in the elderly population is to deal with the problem of depression. It is evident from several studies that there is a steady increase in the incidence and prevalence of depression and related problems among the elderly population. Like in other developed countries, depression in elderly has emerged as a major public health concern in the United States [14]. Birrer and Vemuri [8] observed that nearly 5 million of the 31 million Americans who are 65 years or older are chronically depressed, and 1 million have major depression. The prevalence of depression is higher among elderly females than males. In the primary care settings, 17 to 37 percent have been diagnosed with depression. Recurrence is also very high and can be as high as 40 percent. Birrer and Vemuri [8] further observed that minor depression may follow a major depression and a large proportion of patients with minor depression may develop a major depression within two years [3]. Although minor depression is more common among elderly, they can develop major depressive episodes. Beekman [5] observed from a 6-year longitudinal aging study that the prognosis for depression in late life is poor and the impact of depression is more severe than it was previously thought. The impact of depression at late-life can result in decreased well-being, decreased daily function, increased mortality, and increased use of health services.

It is also noteworthy that the problem of depression is not a natural problem associated with aging but due to the presence of various factors associated with depression or other related problems, the incidence and prevalence of depression problems tend to increase at older age. Kales and Valenstein [30] indicated another problem related to the treatment of depression in elderly population due to complexity in overall management stemming from a number of factors or confounders. Factors such as chronic pain, medical disability, death of spouse etc. can lead a spell of depression in elderly. The presence of accumulated stressors makes it difficult for the elderly population to cope with any episode of depression. It has been documented that there is an association between depression and increased mortality risk in older persons with severe depressive disorder [46]. However, there is a strong belief that the high incidence and prevalence of depression were not measured properly in the past. It was a major concern until recent years that the depression was under recognized in the elderly population but the rates of diagnosis increased rapidly since 1990s.

Crystal *et al.* [13] observed that there are significant disparities by age, ethnicity, and insurance coverage in treatment of the diagnosed elderly population with depression. Fischer *et al.* [15] observed that depression is one of the most prevalent, disabling, and costly chronic health problems seen in primary care. In the USA, the major concerns of primary care physicians working with elderly population include management of depression and dementia [48, 37]. Lundquist *et al.* [37] further observed that these disorders pose the most formidable challenge to the elderly population because these are the major causes of morbidity and mortality in elderly. Various studies [19, 37, 39] observed that 15 percent of the geriatric population is affected by depression. Depression requires greater attention for the elderly population due to high suicide rates among the geriatric patients.

For this study, an application is showed in this section from the Health and Retirement Study (HRS) data. The HRS is sponsored by the National Institute of Aging (grant number NIA U01AG09740) and conducted by the University of Michigan [23]. This study is conducted nationwide for individuals over age 50 and their spouses. We have used the panel data from the six rounds of the study conducted on individuals over age 50 years in 1992, 1994, 1996, 1998, 2000 and 2002. This study uses data documented by RAND. We have used the panel data on depression for the period, 1992-2002. The depression index is based on the score on the basis of the scale proposed by the Center for Epidemiologic Studies Depression (CESD). As indicated in the documentation of the RAND, the CESD score is computed on the basis of eight indicators attributing depression problem. The indicators of depression problem are based on six negative (all or most of the time: depressed, everything is an effort, sleep is restless, felt alone, felt sad, and could not get going) and two positive indicators (felt happy, enjoyed life). These indicators are yes/no responses of the respondents feelings much of the time over the week prior to the interview. The CESD score is the sum of six negative indicators minus two positive indicators. Hence, severity of the emotional health can be measured from the CESD score. From the panels of data, we have used 9761 respondents for analyzing depression among the elderly in the USA during 1992-2002. Steffick [47] indicated that hundreds of studies have used the CESD scale to measure depressive symptoms in a wide range of both clinical and non-clinical populations. Cheng *et al.* [11] showed the validity of a short version of the CESD scale.

We considered the following dependent and explanatory variables: no depression (CESD score $\leqslant 0$) = 0, depression (CESD score > 0) = 1, death =2 if a respondent died during any two consecutive waves; age (in years), gender (male=1, female=0), education (high school or lower=0, some college or higher=1), ethnic group (white=1, else 0; black=1, else 0; others= reference category), number of conditions of health in the past (1-2 condition=1 if the number is 1-2, else 0; 3+ conditions= 1 if the number is 3 or more, else 0; no report of conditions is the reference category), income (low income=1 if the income belongs to lowest 25 percent, else 0, middle income = 1 if belongs to the income group 25 percent to 75 percent, else 0; the highest 25 percent income category is the reference category), duration since the last transition (in years). It is noteworthy that the present study considers only elderly population without any depression symptom at the first wave. In other words, healthy elderly subjects are taken into consideration at the first wave and then depression status is observed for each individual during the subsequent waves.

Table 1 shows counts based on different transition types as displayed in Figure 1. Table 2 summarizes the counts in terms of the proposed models for transitions, reverse transitions and repeated transitions. We considered only subjects with no depression at the baseline and observed that 61 percent of all the subjects remain as depression free during the study period. Out of 8318 depression free subjects, 32 percent moved to the state of depression and 7 percent made transition to death state. Out of the 2660 transitions, 21.1 percent remained in the state of depression during the remaining period (in addition to those who are lost to follow-up during the study period), 72.3 percent made reverse transition to no depression from the immediate prior state of depression and 6.7 percent died from the state of depression. Out of 1922 reverse transitions, 66.4 percent remained in the same state of no depression, 29.3 percent experienced transition to depression once again (repeated transition) and 4.3 percent died from the last prior state of no depression.

Table 3 displays the estimates of the parameters for models on transitions, reverse transitions and repeated transitions as well as for models on transition to death states from no depression, depression (transition), and no depression (reverse). It is observed that age is negatively associated with reverse and repeated transitions. Similarly, gender is negatively associated with all types of transitions (females are more likely to experience all transition types considered here), but after reverse transition, males are more likely to make transition to death state, which is not significant though. We observe that there is no association between education and different types of transition. The white race appears to have significant association with transition to depression but the black race shows positive association to death as compared to other ethnic groups. The interaction between age and gender exert positive impact on transition to depression and reverse transition to depression-free state. The number of health conditions appears to be positively associated with transition to depression, death and repeated transition. The odds of such associations increase with an increased number of conditions. Similar association is observed for low and medium income groups compared to the high income category subjects and higher odds is observed for the lower income group subjects. Duration in a state is expected to have association due to varied intervals between subjects in making transition. It is noteworthy that duration is negatively associated with different types of transitions. Table 4 is constructed excluding the variables education and race which did not show statistically significant association with any transition type or most of the transition types. The models displayed in Tables 3 and 4 are compared using the BIC and the reduced model presented in Table 4 appears to be a better model with lower value of BIC.

From the results, it is evident that age has negative association with repeated transition to depression and with reverse transition as well. Due to lack of any detailed study with all these transition types in the progression of depression, we can not provide much supporting literature. However, in a recent systematic review of 40 studies on the prognosis of depression in older patients in general practice and the community, it is concluded that within the older population, age seems to be a negative prognostic factor [35]. In addition, Freedman *et al.* [17] observed that depression does not increase mortality in elderly. Another study [31] found on the basis of age by cohort interaction that there was evidence of more depression among younger than older women in the post 1944 cohort but a flat age profile in the pre 1945 cohort. Hence, the finding of this study provides some additional insights to explain the ongoing debate in the relationship between age and transition to depression among elderly with additional explanations concerning disease progression at subsequent stages.

This study reveals that transition to depression occurs more among the elderly females as compared to that among males. Cole and Dendukuri [12] concluded on the basis of a systematic review and meta-analysis that female gender appears to be an important risk factor for depression among elderly community subjects. Kasen *et al.* [31] that there is a rise in depression among the women due to shift in family and work-role norms for women.

George and Lynch [18] demonstrated that black population experience greater stress growth than whites and this result in higher depression among blacks. They showed that there is a linear increase in stress growth for blacks but not for whites over time. However, this study does not show any such difference among the elderly population by race.

5. Discussion and conclusion

In this paper, a new approach is proposed to construct the likelihood function for estimation of parameters for models emerging from longitudinal data. The proposed approach takes into account transitions, reverse transitions and repeated transitions and simplifies the construction of likelihood in order to keep the number of models and thereby, number of parameters, to a minimum level without making the underlying theory and analysis complex. It is noteworthy that the proposed method considers any number of follow-ups and any number of transitions for each individual and can be applied to any higher order of underlying Markov chain for modeling data from longitudinal studies. The multistate model proposed in this paper is based on polytomous outcomes which is an extension of repeated binary outcomes. The estimation procedure is illustrated in the paper and it can be seen that the inferential procedure remains simple even after reducing the number of models and parameters by considering the transitions of different types.

The new procedure proposed in this paper is applied to the important stages of transitions that occur among the elderly population in terms of depression. Islam and Chowdhur [27] showed a method for analyzing different stages of binary outcomes with covariate dependence but the proposed technique generalizes it for polytomous outcomes to take account of multiple outcomes at each follow-up. In addition, the method illustrated in this paper shows the procedure to take account of transitions taking place at different follow-ups. This paper provides an extended procedure for dealing with the longitudinal data on depression employing a discrete-time and discrete state Markov model with covariate dependence and shows a more generalized form of the likelihood function in order to model for transitions, reverse transitions and repeated transitions at different follow-ups. In this paper, we have also considered deaths from no depression or depression states as a separate state instead of considering it as a potential source of censoring. We have considered polytomous logistic link functions for analyzing the different types of transitions. The proposed model can be extended easily for continuous-time and discrete-state processes with different link functions.

The problem of depression has emerged as a major health concern during the past decades. This problem can have long-term impact in terms of health risk as well as major socio-economic consequences. In some cases, the transitions, reverse transitions as well as repeated transitions to or from the state of depression can be of interest both to the researchers and to the health professionals. The application of the proposed model to the HRS data on depression for six waves during 1992-2002 reveals some interesting results that reflect the underlying differences in m making transitions, reverse transitions and repeated transitions. The statistical procedure proposed in this paper has the following advantages over the existing techniques: (i) the estimates are obtained for covariate dependence for different models such as transitions, reverse transitions, repeated transitions as well as for transitions to death from no depression (both for first and reverse transitions), and depression (after transition from no depression) from the same global model based on event history data; (ii) it takes into account the conditional nature of transitions in estimating the underlying relationships between covariates and transition probabilities, (iii) the estimates provide a comparable scenario in explaining the nature of relationships for transitions, reverse transitions and repeated transitions and provides a more specific set of results that can explain the underlying relationships more specifically in terms of progression of the disease, etc.

This study provides a generalized approach of covariate dependent Markov chain models for transitions in depression among elderly population in the USA. The results provide important findings associated with transition to depression, reverse transition to no depression and repeated transition to depression that occur over a period of time in the elderly population. The findings of this study provides a more detailed and more comprehensive set of results which can help in explaining the long term patterns in depression as well as the underlying factors that cause transitions to different phases of depression.

Acknowledgement. The HRS (Health and Retirement Study) is sponsored by the National Institute of Aging (grant number NIA U01AG09740) and conducted by the University of Michigan. The authors are grateful to the anonymous reviewers for their useful comments that contributed to improvement in the exposition of the paper to a great extent.

References

[1] P. S. Albert, A Markov model for sequence of ordinal data from a relapsing-remitting disease *Biometrics* **50** (1994), 51–60.
[2] P. S. Albert and M. A. Waclawiw, A two state Markov chain for heterogeneous transitional data: A quasilikelihood approach, *Statistics in Medicine* **17** (1998), 1481–1493.
[3] American Psychiatric Association. (1994). Task force on DSM-IV, diagnostic and statistical manual of mental disorders: DSM-IV, 4th Edition, Washington, DC.: *American Psychiatric Association*.
[4] A. Azzalini, Logistic regression for autocorrelated data with application to repeated measures, *Biometrika* **81** (1994), 767–775.
[5] A. T. Beekman, The natural history of late-life depression. A 6-year prospective study in the community, *Archive Gen Psychiatry* **59** (2002), 605–611.
[6] A. Berchtold and A. E. Raftery, The mixture transition distribution model for high-order Markov chains and non-Gaussian time series, *Statist. Sci.* **17** (2002), no. 3, 328–356.
[7] N. E. Breslow and D. G. Clayton. Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association* **88** (1993), 9–25.
[8] R. B. Birrer and S. P. Vemuri, Depression in later life: A diagnostic and therapeutic challenge, *American Family Physician* **69** (2004), 2375–2386.
[9] P. Bühlmann and A. J. Wyner, Variable length Markov chains, *Ann. Statist.* **27** (1999), no. 2, 480–513.
[10] B. Chen, G. Y. Yi and R. J. Cook, Likelihood analysis of joint marginal and conditional models for longitudinal categorical data, *Canad. J. Statist.* **37** (2009), no. 2, 182–205.
[11] S. T. Cheng, A. C. Chan and H. H. Fung, Factorial structure of a short version of the Center for Epidemiologic Studies Depression Scale, *International Journal of Geriatric Psychiatry* **21** (2006), 333–336.
[12] M. G. Cole and N. Dendukuri, Risk factors for depression among elderly community subjects: A systematic review and meta-analysis, *American Journal of Psychiatry* **160** (2003), 1147–1156.
[13] S. Crystal, U. Sambamoorthi, J. T. Walkup and A. Akincigil, Diagnosis and treatment of depression in the elderly medicare population: Predictors, disparities, and trends, *Journal of American Geriatric Society* **51** (2003), 1718–28.
[14] P. M. Doraiswamy, Contemporary management of of comorbid anxiety and depression in geriatric patients. *Journal of Clinical Psychiatry* **62** (2001), 30–35.
[15] L. R. Fischer, F. Wei, L. I. Solberg, W. A. Rush, and R. L. Heinrich, Treatment of elderly and other adult patients for depression in primary care, *Journal of American Geriatric Society* **51** (2003), 1554–1562.
[16] G. M. Fitzmaurice, N. M. Laird and A. G. Rotnitzky, Regression models for discrete longitudinal responses, *Statist. Sci.* **8** (1993), no. 3, 284–309.
[17] L. Freedman, V. J. Schoenbach, B. H. Kaplan, D. G. Blazer, S. A. James, D. G. Kleinbaum and B. Yankaskas, The association between depressive symptoms and mortality among older participants in the Epidemiologic Catchment Area-Piedmont Health Survey, *Journal of Gerontology* **44** (1989), 149–156.
[18] L. K. George and S. M. Lynch, Race differences in depressive symptoms: A dynamic perspective on stress exposure and vulnerability, *Journal of Health and Social Behavior* **44** (2003), 353–369.

[19] B. Gurland, L. Dean and B. Cross, The epidemiology of depression and dementia in the elderly: The use of multiple indicators of these conditions, In *Psychopathology in the aged.*, O. J. Cole, J. E. Barrett (Eds.), Raven Press, New York, 1980.

[20] J. Harezlak, S. Gao and S. L. Hui, An illness-death stochastic model in the analysis of longitudinal dementia data, *Statistics in Medicine* **22** (2003), 1465–75.

[21] P. J. Heagerty and S. L. Zeger, Marginalized multilevel models and likelihood inference, *Statist. Sci.* **15** (2000), no. 1, 1–26.

[22] P. J. Heagerty, Marginalized transition models and likelihood inference for longitudinal categorical data, *Biometrics* **58** (2002), no. 2, 342–351.

[23] Health And Retirement Study, (Wave [1-6]/Year[1992-2002]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG09740). Ann Arbor, MI.

[24] M. A. Islam, R. I. Chowdhury, N. Chakraborty and W. Bari, A multistage model for maternal morbidity during antenatal, delivery and postpartum periods, *Statistics in Medicine* **23** (2004), 137–158.

[25] M. A. Islam and K. P. Singh, Multistate survival models for partially censored data, *Environmetrics* **3** (1992), 223–234.

[26] M. A. Islam, Multistate survival models for transitions and reverse transitions: an application to contraceptive use data, *J. Roy. Statist. Soc. Ser. A* **157** (1994), no. 3, 441–455.

[27] M. A. Islam and R. I. Chowdhury, A higher order Markov model for analyzing covariate dependence, *Applied Mathematical Modelling* **30** (2006), 477–488.

[28] M. A. Islam, R. I. Chowdhury and L. Briollais, A bivariate binary model for testing dependence in outcomes. *Bulletin of Malaysian Mathematical Sciencies Society (2).* Accepted.

[29] J. D. Kalbfleisch and J. F. Lawless, The analysis of panel data under a Markov assumption, *J. Amer. Statist. Assoc.* **80** (1985), no. 392, 863–871.

[30] H. C. Kales and M. Valenstein, Complexity in late-life depression: Impact of confounding factors on diagnosis, treatment, and outcomes, *Journal of Geriatric Psychiatry Neurology* **15** (2002), 147–55.

[31] S. Kasen, P. Cohen, H. Chen and D. Castille, Depression in adult women: Age changes and cohort effects, *American Journal of Public Health* **93** (2003), 2061–2065.

[32] E. L. Korn and A. S. Whittemore, Methods of analyzing panel studies of acute health effects of air pollution, *Biometrics* **35** (1979), 795–802.

[33] O. Kuss and D. Mclerran, A note on the estimation of the multinomial logistic model with correlated responses in SAS, *Computer Methods and Programs in Biomedicine* **87** (2007), 262–269.

[34] Y. Lee and J. A. Nelder, Conditional and marginal models: Another view, *Statist. Sci.* **19** (2004), no. 2, 219–238.

[35] E. Licht-Strunk, D. A. W. M. Van Der Windt, H. W. J. Van Marwijk, M. De Haan and A. T. F. Beekman, The prognosis of depression in older patients in general practice and the community: A systematic review, *Family Practice* **24** (2007), 168–180.

[36] J. K. Lindsey and P. Lambert, On the appropriateness of marginal models for repeated measurements in clinical trials, *Statistics in Medicine* **17** (1998), 447–469.

[37] R. S. Lundquist, A. Bernens and C. G. Olsen, Comorbid disease in geriatric patients: Dementia and depression – includes patient information, *American Family Physician* **55** (1997), 2687–2694.

[38] L. R. Muenz and L. V. Rubinstein, Markov models for covariate dependence of binary sequences, *Biometrics* **41** (1985), no. 1, 91–101.

[39] S. A. Murrell, S. Himmelfarb and K. Wright, Prevalence of depression and its correlates in older adults, *American Journal of Epidemiology* **117** (1983), 173–185.

[40] G. G. S. Pegram, An autoregressive model for multilag Markov chains, *J. Appl. Probab.* **17** (1980), no. 2, 350–362.

[41] R. Prentice and L. Gloeckler, Regression analysis of grouped survival data with application to breast cancer data, *Biometrics* **34** (1978), 57–67.

[42] A. E. Raftery and S. Tavare, Estimating and modeling repeated patterns in higher order Markov chains with the mixture transition distribution model, *Applied Statistics* **43** (1994), 179–199.

[43] A. E. Raftery, A model for high-order Markov chains, *J. Roy. Statist. Soc. Ser. B* **47** (1985), no. 3, 528–539.

[44] M. H. Regier, A two state Markov model for behavior change, *Journal of the American Statistical Association* **63** (1968), 993–999.

[45] S. M. Ross, *Introduction to Probability Models*, fourth edition, Academic Press, Boston, MA, 1989.

[46] R. A. Schovers, M. I. Geerlings, A. T. Beekman, B. W. Penninx, D. J. Deeg, C. Jonker and W. V. Tilburg, Association of depression and gender with mortality in old age, Results from the Amsterdam study of the elderly, *British Journal of Psychiatry* **177** (2000), 336–342.

[47] D. E. Steffick, *Documentation of Affective Functioning Measures in the Health and Retirement Study.* HRS/AHEAD Documentation Report (DR-005). Survey Research Center, University of Michigan, Ann Arbor, 2000.

[48] M. E. Williams and N. K. Connolly, What practicing physicians in North Carolina rate as their most challenging geriatric medicine concerns? *Journal of American Geriatric Society* **38** (1990), 1230–1234.

[49] M. Wu and J. H. Ware, On the use of repeated measurements in regression analysis with dichotomous responses, *Biometrics* **35** (1979), 513–522.

Table 1. Counts based on transitions, reverse transitions and repeated transition

Transition Types	**N**	**Transition Types**	**N**	**Transition Types**	**N**	**Transition Types**	**N**
No transition		0-0-0-0-0-3	247	0-0-0-1	504	0-0-2	103
0-3	176	0-0-0-0-0-0	3183	0-0-0-0-1	307	0-0-0-2	126
0-0-3	672	Transition 0-1		0-0-0-0-0-1	219	0-0-0-0-2	130
0-0-0-3	463	0-1	1099	Transition 0-2		0-0-0-0-0-2	111
0-0-0-0-3	332	0-0-1	531	0-2	115		
No transition		0-0-0-0-1-3	5	0-0-0-0-1-0	208	Transitions 0-1-2	
0-1-1-1-1-1	113	0-0-0-1-1-3	3	0-0-0-1-1-0	66	0-1-2	50
0-0-1-1-1-1	48	0-0-1-3	7	0-0-1-1-1-0	27	0-0-1-2	17
0-0-0-1-1-1	53	Reverse Tran. 0-1-0		0-1-1-1-1-0	50	0-1-1-2	14
0-0-0-0-1-1	76	0-1-0	246	0-1-0-0-0-0	160	0-0-0-1-2	14
0-0-0-0-0-1	219	0-0-1-0	161	0-0-0-1-0-0	241	0-0-1-1-2	12
0-1-3	16	0-1-0-0	181	0-0-1-0-0-0	162	0-1-1-1-2	15
0-1-1-3	8	0-1-1-0	93	0-0-1-1-0-0	47	0-0-0-0-1-2	18
0-1-1-1-3	5	0-0-0-1-0	109	0-1-1-0-0-0	49	0-0-0-1-1-2	12
0-1-1-1-1-3	2	0-0-1-1-0	41	0-1-1-1-0-0	33	0-0-1-1-1-2	9
0-0-0-1-3	6	0-1-1-1-0	48			0-1-1-1-1-2	16
No transition		0-1-1-0-3	13	0-1-1-0-1-0	16	Transitions 0-1-0-2	
0-0-1-1-0-0	47	0-1-0-0-3	31	0-1-0-0-1-0	38	0-1-0-2	12
0-0-1-0-0-0	162	0-0-1-0-3	32	0-1-0-1	127	0-0-1-0-2	8
0-1-1-1-0-0	33	0-0-0-1-0-3	30	0-0-1-0-1	20	0-1-1-0-2	7
0-0-0-1-0-0	241	0-1-1-1-0-3	10	0-1-1-0-1	23	0-1-0-0-2	17
0-0-0-0-1-0	208	0-0-1-0-0-3	17	0-1-0-0-1	26	0-0-0-1-0-2	10
0-0-0-1-1-0	66	0-1-1-0-0-3	7	0-0-0-1-0-1	69	0-0-1-0-0-2	8
0-1-0-0-0-0	160	0-1-0-0-0-3	18	0-0-1-1-0-1	28	0-1-1-1-0-2	3
0-1-1-0-0-0	49	0-0-1-1-0-3	13	0-1-1-0-0-1	21	0-1-0-0-0-2	11
0-1-1-1-1-0	50	Rep. Tran. 0-1-0-1		0-0-1-0-0-1	37	0-1-1-0-0-2	6
0-0-1-1-1-0	27	0-1-0-1-0-0	45	0-1-1-1-0-1	35		
0-1-0-3	62	0-0-1-0-1-0	39	0-1-0-0-0-1	40		

Note: No depression = 0; Depression = 1; Death = 2; and Missing = 3.

Table 2. Distribution of respondents in terms of the proposed models for transitions, reverse transitions and repeated transitions

Transition Types	N	%
No transition from no depression	5073	61.0
No depression to depression (0→1)	2660	32.0
No depression to death (0→2)	585	7.0
Total	8318	100.0
No transition after transition to depression	561	21.1
No depression to depression to no depression (0→1→0)	1922	72.3
No depression to depression to death (0→1→2)	177	6.7
Total	2660	100.0
No transition after reverse transition	1276	66.4
No depression to depression to no depression to depression (0→1→0→1)	564	29.3
No depression to depression to no Depression to death (0→1→0→2)	82	4.3
Total	1922	100.0

Table 3. Estimates of the parameters of the proposed models for transitions, reverse transitions and repeated transitions

Variables	Transitions									
	0->1					0->2				
	β	SE	OR	95% CI		β	SE	OR	95% CI	
Intercept	0.53	0.71				-7.51	1.13			
Age	0.02	0.01	1.02	1.00	1.05	0.10	0.02	1.10	1.06	1.15
Gender	-3.38	0.84	0.03	0.01	0.18	-2.15	1.20	0.12	0.01	1.23
Education	-0.01	0.07	0.99	0.86	1.14	0.16	0.11	1.17	0.95	1.45
White	0.35	0.16	1.41	1.03	1.94	0.41	0.26	1.50	0.90	2.51
Black	0.29	0.18	1.34	0.95	1.89	0.65	0.27	1.91	1.12	3.27
1-2 Conditions	0.33	0.07	1.39	1.21	1.60	0.76	0.12	2.15	1.70	2.71
3+ Conditions	0.71	0.12	2.03	1.61	2.57	1.80	0.16	6.07	4.48	8.22
Low Income	0.13	0.10	1.14	0.93	1.40	0.76	0.16	2.14	1.56	2.93
Medium Income	0.12	0.09	1.13	0.95	1.35	0.36	0.15	1.43	1.07	1.90
Duration	-0.55	0.02	0.58	0.56	0.60	-0.44	0.02	0.65	0.62	0.68
Age*Gender	0.04	0.01	1.05	1.02	1.08	0.04	0.02	1.04	1.00	1.08
	Reverse Transitions					Transition to Death				
	0->1->0					0->1->2				
Intercept	17.54	1.87				6.68	2.56			
Age	-0.13	0.03	0.88	0.83	0.93	-0.02	0.04	0.98	0.90	1.06
Gender	-7.14	2.85	0.00	0.00	0.21	-3.45	3.58	0.03	0.00	35.40
Education	0.19	0.17	1.21	0.87	1.68	0.32	0.25	1.37	0.85	2.22
White	-0.42	0.41	0.65	0.29	1.47	-0.32	0.54	0.73	0.25	2.12
Black	-0.77	0.44	0.46	0.20	1.09	-0.24	0.57	0.79	0.26	2.39
1-2 Conditions	-0.13	0.17	0.88	0.63	1.21	0.33	0.27	1.40	0.83	2.36
3+ Conditions	-0.28	0.24	0.76	0.47	1.21	1.40	0.32	4.07	2.19	7.58
Low Income	-0.03	0.24	0.97	0.60	1.57	1.17	0.39	3.21	1.49	6.90
Medium Income	0.00	0.22	1.00	0.66	1.53	0.47	0.37	1.61	0.78	3.31
Duration	-1.04	0.05	0.35	0.32	0.39	-1.01	0.06	0.37	0.32	0.41
Age*Gender	0.11	0.05	1.11	1.02	1.21	0.06	0.06	1.06	0.95	1.19
	Repeated Transitions					Reverse Transition and Transition to Death				
	0->1->0->1					0->1->0->2				
Intercept	8.29	1.26				0.64	3.03			
Age	-0.12	0.02	0.88	0.85	0.92	-0.07	0.05	0.94	0.85	1.03
Gender	-0.61	2.18	0.54	0.01	38.71	2.51	4.07	12.33	0.00	35941.94
Education	-0.18	0.13	0.84	0.65	1.09	-0.44	0.34	0.65	0.33	1.25
White	0.15	0.28	1.17	0.68	2.01	-0.73	0.42	0.48	0.21	1.09
Black	0.14	0.30	1.15	0.64	2.08	-0.45	0.46	0.64	0.26	1.55
1-2 Conditions	0.37	0.13	1.45	1.12	1.87	0.79	0.32	2.20	1.18	4.12
3+ Conditions	0.69	0.19	1.99	1.38	2.87	1.46	0.38	4.32	2.05	9.11
Low Income	0.27	0.19	1.32	0.90	1.92	2.46	0.75	11.72	2.68	51.27
Medium Income	0.07	0.17	1.07	0.77	1.51	1.37	0.75	3.92	0.91	16.95
Duration	-0.23	0.02	0.79	0.76	0.83	-0.18	0.04	0.84	0.77	0.91
Age*Gender	0.01	0.04	1.01	0.94	1.08	-0.02	0.07	0.98	0.86	1.11
Model Chi-square	6676.18 (d.f =66, p-value=0.000)									
BIC	14638.31									

Table 4. Estimates of the parameters of the proposed reduced models for transitions, reverse transitions and repeated transitions

Variables	Transitions									
	0->1					0->2				
	β	SE	OR	95% CI		β	SE	OR	95% CI	
Intercept	0.83	0.69				-6.94	1.10			
Age	0.02	0.01	1.02	1.00	1.05	0.10	0.02	1.10	1.06	1.15
Gender	-3.38	0.84	0.03	0.01	0.18	-2.19	1.20	0.11	0.01	1.18
1-2 Conditions	0.33	0.07	1.40	1.22	1.60	0.78	0.12	2.18	1.73	2.74
3+ Conditions	0.71	0.12	2.03	1.61	2.57	1.81	0.15	6.11	4.51	8.26
Low Income	0.12	0.10	1.12	0.93	1.36	0.72	0.15	2.06	1.53	2.78
Medium Income	0.12	0.09	1.13	0.95	1.33	0.32	0.14	1.37	1.04	1.81
Duration	-0.55	0.02	0.58	0.56	0.60	-0.43	0.02	0.65	0.62	0.68
Age*Gender	0.04	0.01	1.05	1.02	1.07	0.04	0.02	1.04	1.00	1.08
	Reverse Transitions					Transition to Death				
	0->1->0					0->1->2				
Intercept	17.24	1.81				6.85	2.48			
Age	-0.13	0.03	0.88	0.83	0.93	-0.03	0.04	0.97	0.90	1.05
Gender	-7.11	2.84	0.00	0.00	0.21	-3.70	3.57	0.03	0.00	27.10
1-2 Conditions	-0.16	0.17	0.86	0.62	1.19	0.32	0.27	1.38	0.82	2.33
3+ Conditions	-0.30	0.24	0.74	0.46	1.18	1.39	0.32	4.03	2.17	7.48
Low Income	-0.20	0.22	0.82	0.53	1.27	1.02	0.37	2.79	1.36	5.69
Medium Income	-0.10	0.21	0.91	0.61	1.35	0.37	0.36	1.45	0.72	2.92
Duration	-1.04	0.05	0.35	0.32	0.39	-1.01	0.06	0.37	0.32	0.41
Age*Gender	0.11	0.05	1.11	1.02	1.21	0.07	0.06	1.07	0.96	1.19
	Repeated Transitions					Reverse Transition and Transition to Death				
	0->1->0->1					0->1->0->2				
Intercept	8.25	1.23				-0.18	3.01			
Age	-0.12	0.02	0.88	0.85	0.92	-0.07	0.05	0.93	0.85	1.03
Gender	-0.58	2.18	0.56	0.01	39.71	2.44	4.07	11.51	0.00	33428.11
1-2 Conditions	0.38	0.13	1.46	1.13	1.88	0.79	0.32	2.21	1.18	4.13
3+ Conditions	0.71	0.19	2.03	1.41	2.92	1.50	0.38	4.50	2.14	9.45
Low Income	0.35	0.18	1.42	1.00	2.01	2.77	0.73	15.99	3.80	67.40
Medium Income	0.13	0.17	1.14	0.82	1.57	1.56	0.74	4.74	1.12	20.13
Duration	-0.23	0.02	0.79	0.76	0.83	-0.18	0.04	0.83	0.77	0.91
Age*Gender	0.01	0.04	1.01	0.94	1.08	-0.02	0.07	0.98	0.86	1.11
Model Chi-square	6643.49 (d.f=48, p-value=0.000)									
BIC	14476.41									