Addiction (1996) 91 (Supplement), S197–S209

SECTION III. METHODS FOR ANALYZING LONGITUDINAL DATA ON RELAPSE

Latent transition analysis for longitudinal data

WAYNE F. VELICER, ROSEMARIE A. MARTIN & LINDA M. COLLINS[1]

Cancer Prevention Research Center, University of Rhode Island, Kingston, Rhode Island & [1]Pennsylvania State University, University Park, Pennsylvania, USA

Abstract

Assessing outcome is a critical problem for the study of addictive behaviors. Traditional approaches often lack power and sensitivity. Latent Transition Analysis is an alternative procedure that is applicable to categorical latent variable models such as stage models. The method involves four different types of parameters, each of which may be relevant to different research questions. Two examples that employ the Stages of Change construct are used to illustrate the method. In the first example, three different models of longitudinal change are compared. In the second example, the effects of an expert system intervention for smoking is compared to a control condition. The method permits the investigation of a series of specific comparisons: (1) the effectiveness of the intervention for individuals in different stages can be assessed; (2) the effectiveness of the intervention can be evaluated for different time intervals; and (3) the effects of intervention on both progression through the stages and regression through the stages or relapse can be assessed. Other potential applications of the method are also discussed.

Introduction

A critical consideration for the study of addictive behaviors is assessing outcome in a longitudinal framework. Traditionally, researchers have depended on criterion measures that have cursory face validity but suffer from some severe limitations including lack of precise definitions, poor statistical power and lack of meaningfulness for some aspects of the problem. These problems with traditional methods of examining outcomes over time may produce results that are misleading. An extensive review of the current measures and a description of some of the problems for the area of smoking cessation is provided by Velicer *et al.* (1992). There are two possible solutions to this problem. The first solution is to develop an alternative measurement model that is multivariate, continuous and sensitive to change across the whole spectrum of the change sequence. Such a model has been proposed by Velicer *et al.* (1996b). The second solution to this problem is to employ alternative methods of statistical analysis. This paper describes a recently developed technique, Latent Transition Analysis (LTA), that represents a sensitive and flexible method for analyzing outcome data where developmental stages are assessed. Examples from the area of smoking cessation will be used to illustrate the advantages of employing latent transition analysis, but the procedure can be applied

Correspondence to: Wayne F. Velicer, PhD, Cancer Prevention Research Center, 2 Chafee Road, University of Rhode Island, Kingston, RI 02881-0808, USA. Tel: +1 (401) 874 2830. Fax: +1 (401) 874 5562.

Submitted 29th March 1996; initial review completed 10th June 1996; final version accepted 2nd August 1996.

Carfax Publishing Company

to any content area where a categorical variable is an appropriate means of assessing outcome.

The Transtheoretical Model (Prochaska & DiClemente, 1983) is a general model for health behavior change. The major organizing dimension of this model is the stage of change (Prochaska & DiClemente, 1983; Velicer *et al.*, 1992). These stages of change have been labeled Precontemplation, Contemplation, Preparation, Action and Maintenance. The definitions of the stages employ both behavior and behavioral intention. The following definitions apply to smoking cessation. *Precontemplation* is a period of time in which smokers are not thinking about quitting smoking within the next 6 months. *Contemplation* is the period of time in which smokers are seriously thinking about quitting smoking in the next 6 months. *Preparation* is a period of time in which smokers are seriously thinking about quitting smoking in the next 30 days and have made at least one quit attempt in the past year. *Action* is a period ranging from 0 to 6 months after smokers have made the overt change of stopping smoking. *Maintenance* is the period beginning 6 months after Action started and continuing until smoking is terminated as a problem. Regression refers to any return from an advanced stage to a previous stage. Relapse refers to regression from Action or Maintenance to a previous stage (DiClemente *et al.*, 1991).

The problems with assessing outcome with a traditional approach can be illustrated from a Stages of Change perspective for the area of smoking cessation. Point prevalence abstinence is the most frequently employed outcome measure in smoking cessation (Velicer *et al.*, 1992). Point prevalence abstinence rates reflect the percentage of people reaching a criterion behavior, usually abstinence, in a specified time frame. Point prevalence for smoking is the percentage of former smokers who are not smoking at a particular point in time. Point prevalence is a discrete outcome measure that results in low statistical power. This lack of power may result in an inability to detect when an intervention is effective and, thus may increase the Type 2 error rate. Point prevalence also lacks meaningfulness. Interventions must facilitate four transitions: (1) Precontemplation to Contemplation, (2) Contemplation to Preparation, (3) Preparation to Action and (4) Action to Maintenance. From a stage of change perspective, point prevalence reflects the percentage of participants taking action at a particular point in time. Typically, only 20% of smokers are in the Preparation stage (Velicer *et al.*, 1995a), and the point prevalence measure, therefore, would be insensitive to progress for the remaining 80% of smokers. It is unlikely that those in early stages will reach criterion behavior and remain abstinent in a short time frame, but effective interventions with early stage smokers will result in changes in point prevalence in an extended follow-up, such as 18 or 24 months post-intervention (Prochaska *et al.*, 1993). Unfortunately, an intervention could be rejected as ineffective when an extended follow-up would have determined that it was effective. Defining outcome in a stage of change manner taps into progress that may not be recognized by the traditional criterion behavior of abstinence. A strong argument exists for measuring outcome from a stage perspective rather than point prevalence (Velicer *et al.*, 1992).

In psychology, the focus of study often is a variable or construct that is not directly observable, such as stage of change. These unobserved constructs organize observed manifest variables and are either static or dynamic (Collins & Cliff, 1990). Latent variables represent general constructs and can be measured with several manifest indicators. Dynamic variables involve systematic change over time while static latent variables are unchanging. Traditional measurement and analysis developed for static variables suffer from serious shortcomings when applied to dynamic variables (Collins & Cliff, 1990). Dynamic latent variables are prominent in the area of substance abuse outcome where change in a longitudinal framework is of interest.

Point prevalence abstinence rates are treated as static variables in analyses. Point prevalence measures are analyzed with repeated analyses for each follow-up time period with each analysis treated independently from the others. From this traditional analysis it is not clear whether those successful at one particular follow-up period are the same people who were successful at the preceding follow-up periods. Only a small minority of contemplators (5%) follow a progressing linear pattern (Prochaska *et al.*, 1991). The multiplicity of analyses may lead to increased Type 1 error whereas the inability to detect point prevalence differences early in follow-ups can lead to increased Type 2 error.

An analysis of stages of change as a dynamic latent variable is able to track individual cases from one time point to another and to clarify who is likely to progress or regress. This is particularly important in substance abuse outcome research where subjects may change outcome status at each follow-up period. Analysis of stage movement would be able to detect intervention effects much earlier. It would also allow analysis of the pattern of change and detection of differential treatment effects for different stages.

The concept of Stage of Change has been successfully applied to a wide variety of areas. Recently, the robust relationship between Stage of Change and Decisional Balance was demonstrated across 12 problem behaviors. Migneault *et al.* (1997) have demonstrated the utility of the stage of change construct for immoderate alcohol use by adolescents. Migneault (1995, unpublished doctoral dissertation) replicated and extended the stage of change construct and other variables of the Transtheoretical Model for immoderate alcohol use with a sample of college students. DiClemente & Hughes (1990) applied the same model to the treatment of alcoholics. Rosenbloom (1990, unpublished doctoral dissertation) applied the model to the area of quitting cocaine. Other problem areas where there is demonstrated empirical support for the utility of the Stages of Change construct include lowering dietary fat (Rossi, 1993, unpublished doctoral dissertation), increasing adherence to mammography-screening schedules (Rakowski *et al.*, 1992), acquisition of regular exercise behaviors (Marcus, Rakowski & Rossi, 1992a; Marcus *et al.*, 1992b), and adoption of condom use for prevention of sexually transmitted diseases (Grimley *et al.*, 1995).

Other prominent stage theories exist within psychology, such as Piaget's stages of cognitive development, Kohlberg's stages of moral development (Kohlberg, 1969) and Kubler-Ross's stages for coping with death (Kubler-Ross, 1975). LTA can be applied to any sequence of discrete states or stages. It employs latent variables, and is most appropriate for a developmental or longitudinal construct when a stage model is the conceptual model guiding longitudinal research.

Latent transition analysis

The most widely employed analysis technique for examining discrete latent variables is *latent class theory* (LCT). LCT is a method for looking at static latent variables that permits estimation of measurement error in the model. In *Latent Class Measurement Theory* (Lazarfeld & Henry, 1968; Dayton & Macready, 1976; Clogg & Goodman, 1984), discontinuous latent variables are measured by observed responses, usually dichotomous, to a manifest indicator variable. Latent class membership is mutually exclusive and each member of a population is classified into one and only one of several latent classes. Latent class theory is limited, however, because it does not handle dynamic latent variables that change systematically over time (Graham *et al.*, 1991). *Markov Models* are a special latent class procedure for stage-sequential dynamic latent variables. Markov procedures are used for predicting the probability of movement through stages over a specific time interval. Markov models are the most widely employed technique for examining discrete dynamic variables longitudinally.

Latent transition analysis (LTA) extends latent class theory and Markov techniques to models that contain both static and dynamic latent variables and an estimation of measurement error. Further, LTA emphasizes the use of multiple indicators allowing testing of complex models.

LTA may be employed to answer several different research questions. First, LTA may be used to test alternative theoretical models about the pattern of change over time. Secondly, LTA is useful for comparing different groups to test for treatment effects. Thirdly, LTA is useful for evaluating the contribution of different measures for each latent status. Finally, LTA identifies the distribution of subjects by latent status at each occasion.

LTA can be performed using a FORTRAN program (Collins, Wugalter & Rousculp, 1991) that uses the Expectation-Maximization (EM) algorithm (Dempster, Laird & Rubin, 1977) for estimating four types of parameters: [1] the gamma parameters (γ), which are estimates of the proportion of the population in each *latent class* (discrete grouping variable); [2] the delta parameters (δ), which are estimates of the proportion of the population in each *latent status* (or stage) at each occasion of measurement, conditional on latent class membership; [3] the tau parameters (τ) that refer to the conditional *probability of making a transition* from one latent status (stage) conditional on previous latent status membership and latent class; and [4] the rho parameters (ρ), which represent *measurement error*, and are estimates of a particular item response conditional on latent status and latent class membership.

The model is most easily understood with an example. Assume there are two occasions of measurement, five latent statuses and a static latent variable with two latent classes. The occasions of measurement will be defined as *time t* for the first and *time* $t+1$ for the second. Latent status represented by S will be defined as PC (precontemplation) for the first latent status, C (contemplation) for the second latent status, P (preparation) for the third latent status, A (action) for the fourth latent status and M (maintenance) for the final latent status. S1 will represent latent status at time 1 and S2 will represent latent status at time 2. Assume that latent status is measured by five items, *item 1*, *item 2*, *item 3*, item 4 and *item 5* where g, h, i, j, k, represent responses at time t and g', h', i', j', k' represent responses at *time* $t+1$. Lastly assume that latent class, i.e. treatment condition (control or intervention), is measured by one item where m equals the response to the item. Therefore, each participant will have a vector of response patterns (Y) for each of the measured variables where $Y=\{m, g, h, i, j, k, g', h', i', j', k''\}$. The formal LTA model can be represented as:

$$\gamma_{\mathrm{LC}} \quad \rho_{\mathrm{M/LC}} \quad \delta_{\mathrm{S1|LC}} \quad \rho_{\mathrm{g|S1,LC}} \quad \rho_{\mathrm{h|S1,LC}} \rho_{\mathrm{i|S1,LC}} \quad \rho_{\mathrm{j|S1,LC}} \quad \rho_{\mathrm{k|S1,LC}} \quad \tau_{\mathrm{S2|S1,LC}} \quad \rho_{\mathrm{g'|S2,LC}} \quad \rho_{\mathrm{h'|S2,LC}} \quad \rho_{\mathrm{i'|S2,LC}} \quad \rho_{\mathrm{j'|S2,LC}} \quad \rho_{\mathrm{k'|S2,LC}}$$

The gamma parameters (γ_{LC}) represent the proportion in each latent class. The gammas sum to one. These do not change across occasions of measurement. LTA models can be run with and without a class variable.

The delta parameters ($\delta_{\mathrm{S1|LC}}$) are the proportion of the population in each of the five latent statuses at each occasion of measurement conditional on latent class. Examination of the delta parameters show the growth or decline in latent status membership over time. Although the LTA procedure computes one delta for each latent status, occasion and latent class combination, only the delta parameters associated with the first occasion must be estimated using the EM algorithm. The remaining delta parameters are functions of other parameters in the model and are computed and printed for the convenience of the user. In this example the LTA procedure produces deltas for each of the five stages of smoking at each of the two occasions for each of the treatment conditions, for a total of 20 parameters. However, only 10 delta parameters, corresponding to the five latent statuses for each of the two treatment conditions at the first time, are estimated.

The tau parameters represent movement across latent statuses over time. Each tau parameter is a probability of being in a particular latent status at Time $t+1$, conditional on latent status membership at Time t and on latent class membership. These parameters are usually arranged in a transition probability matrix where the rows correspond to latent status membership at Time t and the columns correspond to latent status membership at Time $t+1$. This matrix is then row-conditional, i.e. each row sums to one. There will be $C(T-1)$ tau matrices arranged in this fashion, where T is the number of times and C is the number of latent classes.

A schematic representation of a transition probability matrix for the example appears in Table 1. The rows and columns correspond to the five latent statuses in the Stages of Change model. All of the elements in this matrix are conditioned on membership in the Treatment condition. Although it is not shown here, there would be another transition probability matrix for the Control condition in this example. The values on the diagonal of this matrix represent *stability* or the probability of remaining in the same latent status (assuming no one leaves a latent status and returns to it between observations). For example, the element $\mathrm{TP}_{\mathrm{C/PC,T}}$ ($\mathrm{Tp}=$ Tau Parameter) is the probability of membership in the Contemplation stage on the second occasion, conditional on membership in the Precontemplation stage on the first occasion and membership in the Treatment condition. Values above the diagonal of the transition probability matrix represent *progression* or the probability of moving forward to an advanced stage. Values below the diagonal represent *regression* or the probability of moving backward to a previous stage. If according to the model being tested movement among stages can be either forward or backward, all elements of the transition probability matrix are estimated. This is what is shown in Table 1. If according to the model certain kinds of transitions are not possible, the user can fix elements of the transition probability matrix to zero as appropriate. For example, if the two

Table 1. *Full tau parameter matrix for stages of change*

	Stage at occasion 2				
Stage at occasion 1	PC	C	P	A	M
Precontemplation	$\tau_{PC\|PC,LC}$	$\tau_{C\|PC,LC}$	$\tau_{P\|PC,LC}$	$\tau_{A\|PC,LC}$	$\tau_{M\|PC,LC}$
Contemplation	$\tau_{PC\|C,LC}$	$\tau_{C\|C,LC}$	$\tau_{P\|C,LC}$	$\tau_{A\|C,LC}$	$\tau_{M\|C,LC}$
Preparation	$\tau_{PC\|P,LC}$	$\tau_{C\|P,LC}$	$\tau_{P\|P,LC}$	$\tau_{A\|P,LC}$	$\tau_{M\|P,LC}$
Action	$\tau_{PC\|A,LC}$	$\tau_{C\|A,LC}$	$\tau_{P\|A,LC}$	$\tau_{A\|A,LC}$	$\tau_{M\|A,LC}$
Maintenance	$\tau_{PC\|M,LC}$	$\tau_{C\|M,LC}$	$\tau_{P\|M,LC}$	$\tau_{A\|M,LC}$	$\tau_{M\|M,LC}$

occasions of measurement are only 4 months apart, it is not possible to move from the contemplation stage to the maintenance stage, because the action stage between them is defined as a six month period.

There are two sets of rho parameters (ρ); one associated with the static latent variable representing latent class membership and one associated with the dynamic latent variable representing latent status membership. The rhos (ρ) represent the probability of a particular response to each manifest variable at each occasion of measurement conditional on latent class membership and/or latent status membership. The measurement parameters ($\rho_{M|LC}$) represent the probability of responses to the item measuring the static latent variable conditional on latent class membership. In other words, what is the probability a member of latent class one will select the first response category. The rho parameters associated with the dynamic latent variable represent the probability of response to the item measuring the dynamic latent variable conditional on latent status membership and latent class membership. LTA may be used to test either latent or manifest variable problems. If there is only one measure for latent status membership, the model tested is a manifest variable problem. Latent variable problems require two or more manifest variables measuring each latent status.

The ρs serve two roles in the LTA model. The ρs map the manifest item onto the latent statuses in much the same way as factor loadings show the relationship between items and a factor. The ρs also show how precisely the manifest items measure the latent variables. The rho parameters can be interpreted as representing the relationship between the manifest variables and latent classes in the same way the factor loadings relate the manifest variables to the latent factors in structural equation modeling. However, because rho parameters are probabilities rather than regression weights, they are scaled differently from factor loadings. The values close to 0 or 1 indicate that the manifest response is determined by latent status membership and values that are close to one divided by the number of response patterns are determined by chance.

LTA is used to determine how well a particular model fits the data. As the mathematical representation of the model shows, the model predicts the number of people who will contribute to a particular response pattern. A good fitting model will have predictions that are close to the observed number of people who have contributed each response pattern in the actual data. The goodness of fit statistic will be roughly equal to degrees of freedom for a good fitting model (Graham *et al.*, 1991). G^2, a likelihood-ratio goodness-of-fit statistic, is approximately distributed as a χ^2 where degrees of freedom is number of possible response patterns minus number of estimated parameters minus one or expressed $df = K - P - 1$. A necessary but not sufficient condition for identification is positive degrees of freedom, i.e. the number of parameters to be estimated does not exceed the number of response patterns minus one. However, in many instances this condition is met and there are nevertheless identification problems. Underidentification occurs when there is not enough information in the data to estimate all the parameters the user would like to estimate. The problem can be addressed by reducing the size of the model by means of constraining or fixing parameters. When a group of parameters is constrained to be equal to each other, the entire group counts as one estimated parameter, reducing the number of parameters to be estimated. When a parameter is fixed to a prespecified

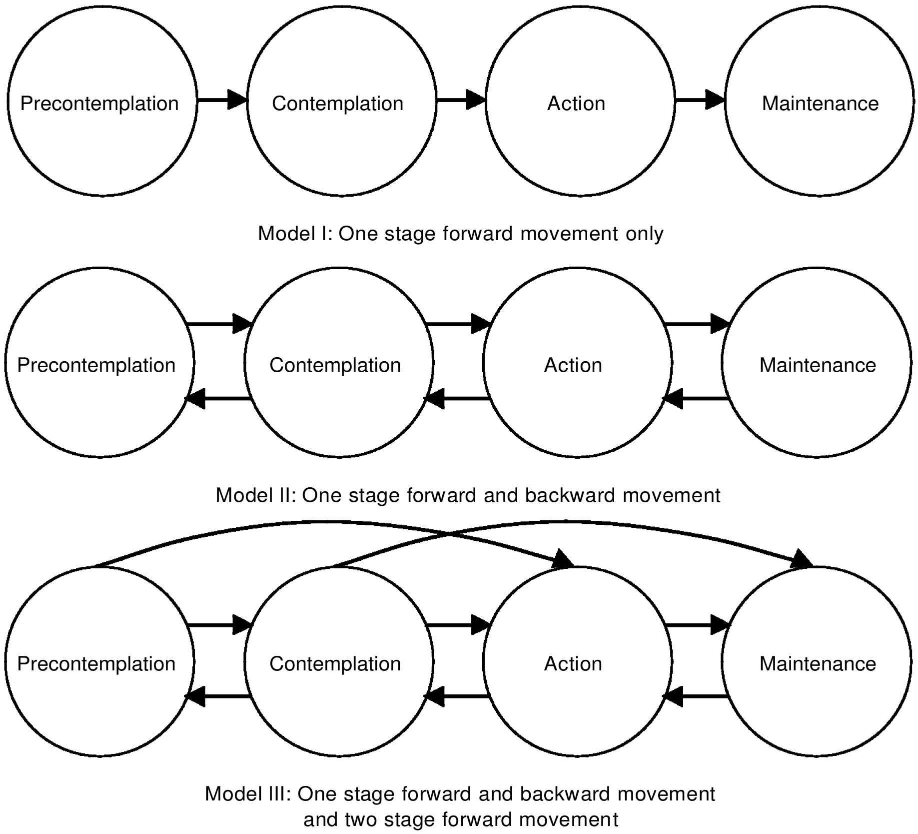


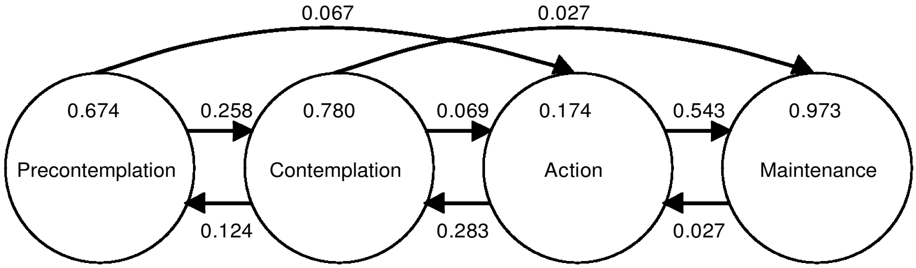
Figure 1. *Three alternative models to be compared by Latent Transition Analysis.*

value, as when a tau parameter is fixed as zero, no estimation is performed for that parameter.

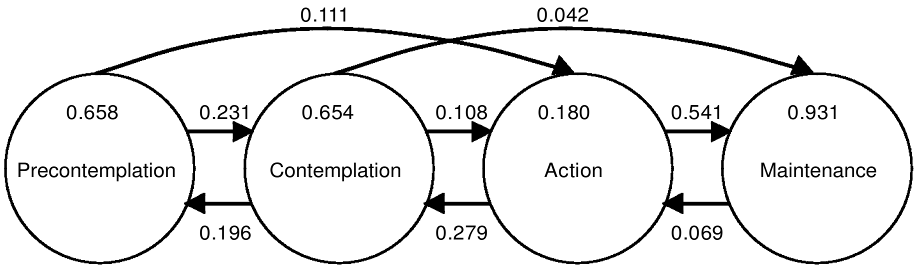
LTA as a test of theoretical models

To illustrate the use of LTA to test alternative theoretical models, an example of LTA applied to the Stages of Change for smoking cessation will be described (Martin, Velicer & Fava, 1996). A sample of 545 reactively recruited smokers or former smokers were assessed five times over a 2-year period, with assessments at 6-month intervals. The smokers were recruited through a series of newspaper advertisements. No intervention was applied. An algorithm was applied based on a series of questions to classify subjects into four stages: Precontemplation, Contemplation, Action and Maintenance (Prochaska & DiClemente, 1983). (At the time this study began, Preparation had not been conceptualized as a separate stage.) LTA was used to determine which of several alternative models would best explain movement between the stages and to describe the pattern of transitions between the stages.

The four stages represent four latent statuses and, at any time, a smoker is in one of these latent statuses. Models are specified by fixing, constraining, or freely estimating parameter values. Since there was no intervention in this study, there was no latent class variable in this example. Figure 1 illustrates three alternative models that were tested and compared. Model I proposed only one stage forward movement among the four latent statuses. Model II proposed both forward and backward movement, but limited to one stage between occasions. Model III proposed both forward and backward movement, and allowed both one stage forward and backward and two stage forward movement. Models were specified by fixing the appropriate transition probabilities to zero.



Tau parameter estimates for Model III, Time 1 to 2



Tau parameter estimates for Model III, Time 2 to 3

Figure 2. *Tau parameter estimates for Model III for the Time 1 to Time 2 transition and Time 2 to Time 3 transition (reprinted with permission from Martin, Velicer & Fava, 1996).*

Since Model I permitted only one stage forward movement all transitions backward, or below the matrix diagonal, would be fixed at zero along with all transitions in the far off diagonals. Similarly, since Model II permitted only one stage forward and backward movement, all transitions more than one step removed from the diagonal would be fixed as zero.

A Goodness of Fit statistic, G^2, approximately distributed as a χ^2 is used to assess model fit and compare alternative models (Graham *et al.*, 1991). Nested models were compared by G^2 difference tests. Model III, the most general model, was found to fit the data best. This model indicates that both one stage progressions and regressions among the stages occur as well as two stage progressions. Figure 2 presents the transition probabilities for the first two transitions (baseline to 6-month follow-up and 6-month follow-up to 12-month follow-up). Values within each circle represent the subjects who remain in each stage, values with the forward arrows represent progression, and values with the backward arrows represent regression. Since these are conditional probabilities, the values inside the circle and values leading away from the circle sum to 1.00. The probability of progressing was greater than that of regressing. The probability of movement to the adjacent stage was greater than that of progressing two stages (for additional details, see Martins *et al.*, 1996). Tests of significance associated with the individual tau parameters were not available in the version of the program used for this analysis.

LTA for comparing treatment and control

The use of LTA for evaluating treatment effects will be illustrated using a second example. The effects of an expert system intervention (Velicer *et al.*, 1993) was compared to an assessment only group with a large representative sample of 3840 smokers (Fava, Velicer & Prochaska, 1994). Subjects were proactively recruited, i.e. contacted by a random digit dial phone call and offered the opportunity to participate after screening for eligibility. The sample consisted only of smokers at baseline. Subjects were assessed on five occasions over a 2-year period, with assessments at 6-month intervals. An algorithm was used to classify subjects into five stages of change: Precontemplation, Contemplation, Preparation, Action and Maintenance. Only cur-

Table 2. *Delta parameter estimates (probabilities of latent status membership) for the intervention group (reproduced with permission from Velicer* et al., *1996a)*

	Occasion				
Stage	1	2	3	4	5
Precontemplation	0.36	0.28	0.22	0.17	0.14
Contemplation	0.37	0.32	0.27	0.24	0.21
Preparation	0.27	0.26	0.27	0.26	0.26
Action	0.00	0.15	0.20	0.23	0.25
Maintenance	0.00	0.00	0.05	0.10	0.14

rent smokers were eligible to participate. Therefore, only the Precontemplation, Contemplation and Preparation stages were represented at baseline. The intervention is a series of 3–4-page completely individualized reports which provide information about how to change the problem behavior. The reports were generated by the expert system and were mailed immediately following assessment. There were three reports: at baseline, 3 months and 6 months. This intervention is described in detail elsewhere (Velicer *et al.*, 1993, 1994) as is a complete description of the LTA analysis for this study (Velicer *et al.*, 1996a).

Since latent variable models should be invariant across samples, the three models described earlier were again compared on this second sample. Model III best represented the data as in the previous example. The analysis estimates four types of parameters for each occasion of measurement to the next occasion.

The gamma parameters (γ) represents the proportion in each latent class. The latent class in this example is the experimental manipulation that resulted in two latent classes. The gamma parameters represent the proportion in the expert system intervention group and the control group. Latent class could also be a naturally occurring characteristic such as gender and may have more than two classes. Latent class is invariant across occasions of measurement.

The delta parameters (δ) are the proportion in each stage at each occasion of measurement, conditional on membership in the intervention groups (latent class). There is one delta for each stage at each occasion of measurement and latent class, although only the deltas for the first occasion are estimated. An example of delta parameters for the intervention group are presented in Table 2.

Examination of the delta parameters shows the changes in distribution stage membership over time. The Precontemplation and Contemplation stages membership gradually declines over the five occasions. The Preparation stage has the same proportion across the occasions. The Action stage shows zero percent membership since only smokers were eligible for the study. At the second occasion of measurement 6 months later, the Action stage shows a steady increase in membership. The Maintenance stage follows a similar pattern to that of Action.

An example of a transition probability matrix for the expert system intervention group is presented in Table 3. The probabilities are conditional on stage membership and latent class for the Time 3 to Time 4 transition when all stage memberships are represented. The diagonal elements of the tau matrix represent stability, i.e. the proportion of individuals who remain in the same stage on both occasions and are displayed in Part I of Table 3. The elements above the diagonal represent progression and are displayed in Part II of Table 3. These values represent the proportion of individuals who move forward to a different stage on the second occasion. The elements below the diagonal represent regression, i.e. the proportion of individuals who move backward to a previous stage on the second occasion and are displayed in Part III of Table 3. Comparing the above diagonal values and below diagonal values provides an overall summary of the effectiveness of an intervention.

The tau parameter values for two latent classes can be compared for differential treatment effects. Figure 3 presents graphically the transition

Table 3. *Tau parameter estimates for the expert system intervention group, probabilities of latent status transitions, for Model III (reproduced with permission from Velicer* et al., *1996a)*

Stage at Time 3	Stage at Time 4 PC	C	P	A	M
Part I. Stable cells					
Precontemplation	0.49	0.29	0.22	0.00	0.00
Contemplation	0.24	0.33	0.22	0.21	0.00
Preparation	0.00	0.33	0.36	0.32	0.00
Action	0.00	0.00	0.30	0.34	0.36
Maintenance	0.00	0.00	0.00	0.47	0.53
Part II. Progressing cells					
Precontemplation	0.49	0.29	0.22	0.00	0.00
Contemplation	0.24	0.33	0.22	0.21	0.00
Preparation	0.00	0.33	0.36	0.32	0.00
Action	0.00	0.00	0.30	0.34	0.36
Maintenance	0.00	0.00	0.00	0.47	0.53
Part III. Regressing cells					
Precontemplation	0.49	0.29	0.22	0.00	0.00
Contemplation	0.24	0.33	0.22	0.21	0.00
Preparation	0.00	0.33	0.36	0.32	0.00
Action	0.00	0.00	0.30	0.34	0.36
Maintenance	0.00	0.00	0.00	0.47	0.53

probabilities for both the expert system intervention group and control group.

The proportion of subjects progressing within each group can be compared. Generally, the intervention group has a larger proportion of subjects progressing when compared to the control group. Fifty-one per cent of Precontemplators in the intervention group progress to Contemplation (29%) or Preparation (22%) where only 41% of Precontemplators in the control group progress. The Contemplation and Preparation stages also follow this same pattern. Of the subjects in the intervention group who were in Preparation at Time 3, 32% progressed to Action, whereas only 27% progressed in the control group. The proportion of subjects progressing out of Action to Maintenance was about the same for both groups.

The proportion of subjects regressing can be compared in the same fashion. The values for the proportion regressing are essentially the same for both groups for the Time 3 to Time 4 transition. Only the Preparation stage differs between the two groups, where 33% of the intervention group regresses compared with 37% of the control group.

Similarly, we can examine if the proportion of stable subjects is different for treatment than for

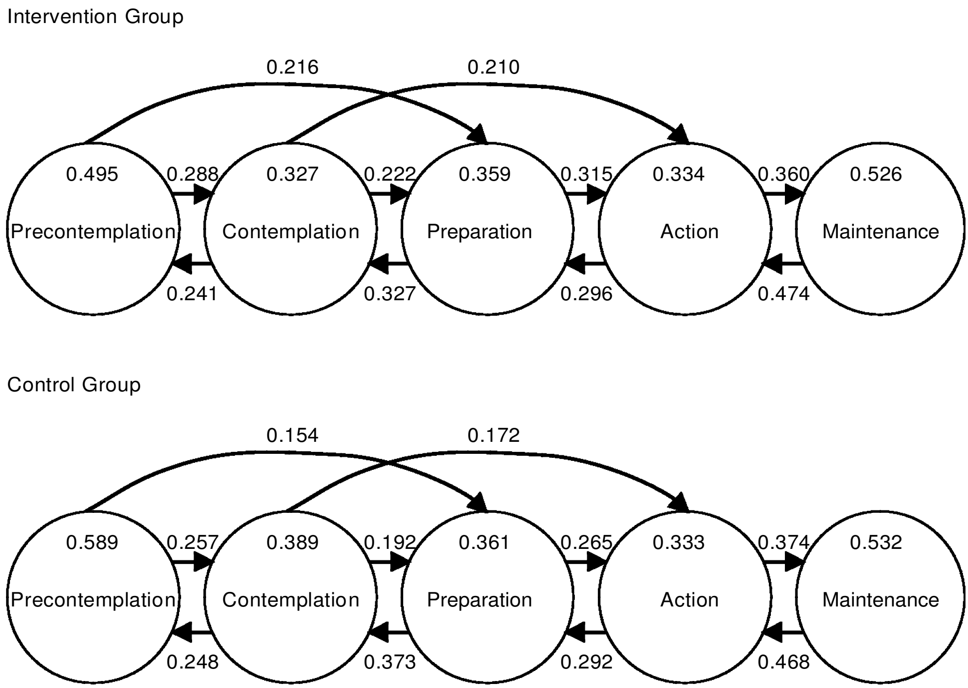


Figure 3. *Tau parameter estimates for Model III for Time 3 to Time 4 for the expert system intervention group and the control group (reproduced with permission from Velicer* et al.*, 1996).*

control. Fifty per cent of Precontemplators in the intervention group remain in Precontemplation and an even greater proportion of the control group (59%) remain. Comparison of the Contemplation stage has the same pattern. The proportion remaining in Preparation, Action and Maintenance are about equal for both the intervention and control groups.

These comparisons employ the total sample resulting in a more powerful analysis. Consider again the point prevalence measure most often used to evaluate outcome. As a dichotomous variable, the only group comparison made can be between the proportion of subjects smoking versus not smoking. In the LTA framework much more information is provided. The intervention group indeed has more forward movement and less stability and regression than the control group for the early stages of Precontemplation and Contemplation. For the later stages this contrast is not evident. This is an important finding since with point prevalence the treatment effects found with the early stages would be missed completely. This intervention may have been rejected when it is effective.

The treatment–control differences may change when early transitions are compared with later transitions. (In this case, treatment occurred between Time 1 and Time 2 and between Time 2 and Time 3. The numeric example uses data for a transition after treatment was completed.) Treatment effects immediately post-intervention may be quite different from long-term effects. LTA permits comparison of transition probability patterns for different transitions. Smokers in the early stages of Precontemplation and Contemplation may take some time to change their smoking status. However, the successful interventions with early stage smokers seen in this example will result in changes in point prevalence in an extended follow-up, such as 18 or 24 months post-intervention (Prochaska *et al.*, 1993). Once again, this intervention may have been rejected when it is effective at an extended follow-up.

In LTA models all transitions contribute, resulting in a more meaningful outcome analysis. All probabilities are conditional on being in a stage on the previous occasion and the sum of these probabilities must be 1.00. Each stage can

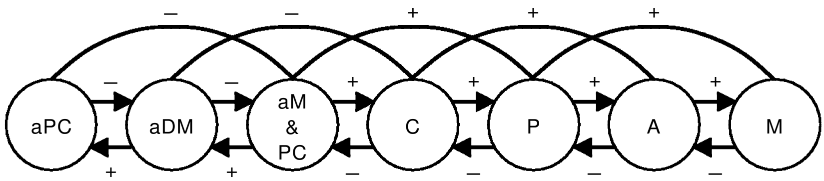


Figure 4. *Stages of change for acquisition and cessation in an integrated LTA model (− = reduce; + = increase).*

be evaluated separately as in this example. But all transitions are meaningful. Two traditional measures, point prevalence and sustained abstinence are the same as the last two transitions. The proportion of the sample in either Action or Maintenance (i.e. the non-smokers) is the same as point prevalence. The proportion in Maintenance is the same as sustained abstinence (i.e. at least 6 months not smoking). Therefore, LTA includes the traditional outcome measures as special cases of a more general model.

Discussion

There are some important considerations when applying LTA. First, an important limitation is that LTA requires large sample sizes. We can consider each possible transition as a separate contingency table. Often this table will contain a large number of possible response patterns. Indeed, many of the cells that have been sampled may be empty. Under these circumstances, the G^2 distribution is not well approximated by a χ^2. However, the distribution of the G^2 difference that is used for nested models is more robust. The larger the sample size, the more likely the contingency table cells will be less sparse. Other limitations at this time include the general lack of available software (only one program is currently available), long computational times for the analysis, and tests of significance that are not fully developed. With regard to the latter, it is important to note that null hypothesis testing is compatible with LTA and will be a feature of the next iteration of LTA software which should be available within a year. Linda M. Collins can provide a copy of the current LTA program and information about new upcoming versions. See Acknowledgements for her address.

On the other hand, LTA is an extremely flexible method. There is a wide variety of potential applications for this method of data analysis and the examples discussed here represent only a limited sampling. To illustrate some of the potential applications, we will describe three other potential analyses.

The rho parameters (ρ) in the LTA model are associated with both latent *class* membership and latent *status* membership. These parameters represent the probability of a particular response to each manifest variable on each occasion of measurement conditional on latent class membership and latent status membership. The two examples presented in this paper did not focus on rho parameters. Only one measure of stage was employed and verification measures were available to correct erroneous stage classification. Therefore, the two analyses presented here are examples of manifest variable problems. If two or more manifest variables are available to determine stage membership, the analysis would be a latent variable problem. In a future study, the classification algorithm used in these examples could be compared to alternate measures of stage membership such as the University of Rhode Island Change Assessment (URICA; McConnaughy, Prochaska & Velicer, 1983; McConnaughy *et al.*, 1989) and the contemplation ladder (Biener & Abrams, 1991) using the rho parameters.

All examples presented here involved adult data where the focus is on the Stages of Change for cessation. The study of addictive behaviors with adolescents includes both acquisition and cessation. For smoking cessation, the Stages of Acquisition were originally identified by Stern *et al.*, 1987. A similar combination of acquisition and cessation has been identified for immodest drinking by adolescents (Migneault *et al.*, 1997). There are fewer stages of acquisition than there are for cessation. The middle stages seem to be a combination of cognition and exploratory actions, and has been labeled as *acquisition Decision Making* (aDM). The two end stages are *acquisition Precontemplation* and *acquisition Maintenance*. This latter stage is the same as Precontemplation for Cessation. Adequately representing adolescent addictive behavior requires a model that includes both acquisition and cessation. Figure 4 illustrates this.

The goals of an intervention program are to reduce progression and increase regression in the acquisition stages and to increase progression and decrease regression in the cessation stages. Graphically, in Fig. 4, when '+' applies to forward paths and '−' to backward paths, a prevention model exists. When '+' applies to forward paths and '−' to backward, a cessation model exists. This illustrates the necessity for adolescent intervention programs to be individualized to the needs and current stage of each person rather than attempting to implement a single intervention for all adolescents.

As an example, Migneault *et al.* (1997) applied the stage of change model for cessation and acquisition to the domain of immoderate alcohol use among adolescents. Of the 853 secondary school students in the sample, 519 were in the stages of acquisition. Precontemplation for ac-quisition (aPC) is the period of time when there is no intention to start immoderate drinking in the next 6 months. Contemplation, Action and Preparation are combined into a Decision Making for acquisition (aDM) stage when there is intention to start immoderate drinking in the near future. Maintenance for acquisition (aM) is a period when there is immoderate drinking and no intention to quit. Maintenance for acquisition is equivalent to Precontemplation for cessation.

The focus of the illustrations presented in this paper have been on the tau values that represent progress compared to the tau parameters that represent regression. However, the diagonal or stable values (see Table 3) also contain important information. For example, one of the important questions about the Stages of Change for both theoretical and practical reasons is the relative stability of the five stages. Although neither of the examples cited here are ideal for answering this question—the first did not separate Contemplation and Preparation and the second study included only the first three stages in the initial recruitment—the combination of the two studies can provide some tentative answers to this question. The two end stages, maintenance and precontemplation should be the most stable (Velicer *et al.*, 1996b). In Fig. 2, the extremely high values for maintenance (0.973 and 0.931) illustrate the stability of this stage. Precontemplation is also very stable, but less so in the intervention study. The middle three stages should be the least stable, and they are (see Fig. 3). However, Contemplation should be more stable than either Preparation or Action (Velicer *et al.*, 1996b) and this is not the case (see Fig. 3), perhaps because of the presence of an intervention.

The topics discussed in this paper illustrate some of the advantages of LTA over traditional outcome analysis. LTA estimates four types of parameters for each specified model. These parameter estimates provide extremely valuable information for comparing theoretical models, evaluating differences between groups and evaluating the contribution of different manifest variables to the measurement of the latent status. LTA models represent an alternative to traditional outcome analysis that is more powerful, more meaningful, and more sensitive.

Acknowledgements

This research was partially supported by Grants CA27821 and CA50087 from the National Cancer Institute and DA04111 from the National Institute on Drug Abuse. An early version of this paper was presented by the first author at the NIAAA Workshop, Methodologies for Relapse Research, Santa Fe, New Mexico, USA, January, 1993. The LTA Program is available from Linda M. Collins, The Methodology Center and Department of Human Development and Family Studies, S-159 Henderson Building South, The Pennsylvania State University, University Park PA 16802-6504, USA.

References

BIENER, L. & ABRAMS, D. B. (1991) The contemplation ladder: validation of a measure of readiness to consider smoking cessation, *Health Psychology*, 10, 360–365.

CLOGG, C. C. & GOODMAN, L. A. (1984) Latent structure analysis of a set of multidimensional contingency tables, *Journal of the American Statistical Association*, 79, 762–771.

COLLINS, L. M. & CLIFF, N. (1990) Using the longitudinal Guttman simplex as a basis for measuring growth, *Psychological Bulletin*, 108, 128–134.

COLLINS, L. M. & WUGALTER, S. E. (1992) Latent class models for stage-sequential dynamic latent variables, *Multivariate Behavioral Research*, 27, 131–137.

COLLINS, L. M., WUGALTER, S. E & ROUSCULP, S. S. (1991) *LTA User's Manual* (Los Angeles, J. P. Guilford Laboratory of Quantitative Psychology, University of Southern California).

DAYTON, C. M. & MACREADY, G. B. (1976) A probabilistic model for validation of behavioral hierarchies, *Psychometrika*, 41, 189–204.

DEMPSTER, A. P., LAIRD, N. M. & RUBIN, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society*, 39, 1–38.

DICLEMENTE, C. C. & HUGHES, S. O. (1990) Stages of Change Profiles in Outpatient alcoholism treatment, *Journal of Substance Abuse*, 2, 217–235.

DICLEMENTE, C. C. & PROCHASKA, J. O. (1982) Self-change and therapy change of smoking behavior: a comparison of processes of change of cessation and maintenance, *Addictive Behaviors*, 7, 133–142.

DICLEMENTE, C. C., PROCHASKA, J. O., FAIRHURST, S. K., VELICER, W. F., VELASQUEZ, M. M. & ROSSI, J. S. (1991) The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change, *Journal of Consulting and Clinical Psychology*, 59, 295–304.

FAVA, J. L., VELICER, W. F. & PROCHASKA, J. O. (1994) Applying the Transtheoretical Model to a large representative smoking sample, *Addictive Behaviors*, 19, 189–203.

GRAHAM, J. W., COLLINS, L. M., WUGALTER, S. E., CHUNG, N. J. & HANSEN, N. B. (1991) Modeling transitions in latent stage-sequential processes: a substance use prevention example, *Journal of Consulting and Clinical Psychology*, 59, 48–57.

GRIMLEY, D. M., PROCHASKA, J. O., VELICER, W. F. & PROCHASKA, G. E. (1995) Contraceptive and condom use adoption and maintenance: a stage paradigm approach, *Health Education Quarterly*, 22, 18–33.

KOHLBERG, L. (1969) Stage and sequence: the cognitive–developmental approach to socialization, in: GOSLIN, D. A. (Ed.) *Handbook of Socialization Theory and Research* (Chicago, Rand McNally).

KUBLER-ROSS, E. (1975) *Death: The Final Stage of Growth* (Englewood Cliffs, NJ, Prentice-Hall).

LAZARFELD, P. F. & HENRY, N. W. (1968) *Latent Structure Analysis* (Boston, Houghton Mifflin).

MARCUS, B. H., RAKOWSKI, W. & ROSSI, J. S. (1992a) Assessing motivational readiness and decision-making for exercise, *Health Psychology*, 11, 257–261.

MARCUS B. H., ROSSI, J. S., SELBY, V. C., NIAURA, R. S. & ABRAMS, D. B. (1992b) The stages and processes of exercise adoption and maintenance in a worksite sample, *Health Psychology*, 11, 257–261.

MARTIN, R. A., VELICER, W. F. & FAVA, J. L. (1996) Latent transition analysis applied to the stages of change for smoking cessation, *Addictive Behaviors*, 21, 67–80.

MCCONNAUGHY, E. A., DICLEMENTE, C. C., PROCHASKA, J. O. & VELICER, W. F. (1989) Stages of change in psychotherapy: a follow-up report, *Psychotherapy*, 26, 494–503.

MCCONNAUGHY, E. A., PROCHASKA, J. O. & VELICER, W. F. (1983) Stages of change in psychotherapy: measurement and sample profiles, *Psychotherapy: Theory, Research, and Practice*, 20, 368–375.

MIGNEAULT, J. P., PALLONEN, U. E. & VELICER, W. F. (1997) Decisional balance and stage of change for adolescent drinking, *Addictive Behaviors*, 22, 1–13.

PROCHASKA, J. O. & DICLEMENTE, C. C. (1983) Stages and processes of self-change of smoking: toward an integrative model of change, *Journal of Consulting and Clinical Psychology*, 51, 390–395.

PROCHASKA, J. O., DICLEMENTE, C. C., VELICER, W. F. & ROSSI, J. S. (1993) Standardized, individualized, interactive, and personalized self-help program for smoking cessation, *Health Psychology*, 12, 399–405.

PROCHASKA, J. O., VELICER, W. F., DICLEMENTE, C. C. & FAVA, J. L. (1988) Measuring processes of change: applications to the cessation of smoking, *Journal of Consulting and Clinical Psychology*, 56, 520–528.

PROCHASKA, J. O., VELICER, W. F., DICLEMENTE, C. C., GUADAGNOLI, E. & ROSSI, J. S. (1991) Patterns of change: a dynamic typology applied to smoking cessation, *Multivariate Behavioral Research*, 24, 83–107.

PROCHASKA, J. O., VELICER, W. F., ROSSI, J. S. *et al.* (1994) Stages of Change and Decisional Balance for Twelve Problem Behaviors, *Health Psychology*, 13, 39–46.

RAKOWSKI, W., DUBE, C., MARCUS, B. H., PROCHASKA, J. O., VELICER, W. F. & ABRAMS, D. B. (1992) Assessing elements of women's decision about mammography, *Health Psychology*, 11, 111–118.

STERN, R. A., PROCHASKA, J. O., VELICER, W. F. & ELDER, J. P. (1987) Stages of adolescent cigarette smoking acquisition: measurement and sample profiles, *Addictive Behaviors*, 12, 319–329.

VELICER, W. F., DICLEMENTE, C. C., PROCHASKA, J. O. & BRANDENBURG, N. (1985) Decisional balance measure for assessing and predicting smoking status, *Journal of Personality and Social Psychology*, 48, 1279–1289.

VELICER, W. F., DICLEMENTE, C. C., ROSSI, J. S. & PROCHASKA, J. O. (1990) Relapse situations and self-efficacy: an integrative model, *Addictive Behaviors*, 15, 271–283.

VELICER, W. F., PROCHASKA, J. O., BELLIS, J. M. *et al.* (1993) An expert system intervention for smoking cessation, *Addictive Behaviors*, 18, 269–290.

VELICER, W. F., FAVA, J. L., PROCHASKA, J. O., ABRAMS, D. B., EMMONS, K. M. & PIERCE, J. P. (1995a) Distribution of smokers by stage in three representative samples, *Preventive Medicine*, 24, 401–411.

VELICER, W. F., HUGHES, S. L., FAVA, J. L., PROCHASKA, J. O. & DICLEMENTE, C. C. (1995b) An empirical typology of subjects within stage of change, *Addictive Behaviors*, 20, 299–320.

VELICER, W. F., MARTIN, R. A., FAVA, J. L. & PROCHASKA, J. O. (1996a) *Latent Transition Analysis: A Sensitive Outcome Analysis Method* (Kingston, RI, Cancer Prevention Research Center Paper, University of Rhode Island).

VELICER, W. F., PROCHASKA, J. O., ROSSI, J. S. & SNOW, M. G. (1992) Assessing outcome in smoking cessation studies, *Psychological Bulletin*, 111, 23–41.

VELICER, W. F., ROSSI, J. S., PROCHASKA, J. O. & DICLEMENTE, C. C. (1996b) A criterion measurement model for health behavior change, *Addictive Behaviors*, 21, 1–30.

VELICER, W. F., ROSSI, J. S., RUGGIERO, L. & PROCHASKA, J. O. (1994) Minimal interventions appropriate for an entire population of smokers, in: RICHMOND, R. (Ed.) *Interventions for Smokers: An International Perspective*, pp. 69–92 (Baltimore, Williams & Wilkins).

WIGGINS, L. M. (1973) *Panel Analysis* (Amsterdam, Elsevier).