Prevention Science, Vol. 5, No. 3, September 2004 (© 2004)

Defining Subgroups of Adolescents at Risk for Experimental and Regular Smoking

Lisa C. Dierker,[1,5] Shelli Avenevoli,[2] Abbie Goldberg,[3] and Meyer Glantz[4]

If multiple etiologies of substance use are truly at work in the population, then further strides in the accurate prediction of smoking and the use of other substances will likely be built on diverse pattern-centered approaches that explore the presence of multiple population subgroups across various substance use stages. The present study aimed to identify population subgroups defined by individual risk factors or risk factor constellations that prospectively predict specific smoking stages. Using data from the National Longitudinal Study of Adolescent Health (Add Health), analyses were conducted on the sample that took part in the baseline and 1 year follow-up assessment between 1994 and 1996. Classification and regression tree procedures were used to investigate the structure of individual risk factors, or constellations of risk, that define population subgroups with high rates of both experimental and established smoking. For each level of smoking, a relatively simple model including two subgroups predicted over half of the smoking cases. Findings also indicated that the two group models identified higher rates of regular smokers compared to experimental smokers. Deviant behaviors and alcohol use without permission independently predicted movement to experimentation at follow-up. Progression to regular smoking from both a nonsmoking and experimental smoking status at baseline were each predicted by smoking friends. Additionally, baseline levels of experimental use predicted movement from experimental to regular smoking, while a relatively low grade point average predicted rapid progression from baseline nonuse to regular use at follow-up. By identifying first approximations of patterns, these analyses may lead to clues regarding the major multiple mechanisms at work for the progression of smoking among adolescents.

KEY WORDS: smoking; risk factors; substance use; adolescents.

INTRODUCTION

The goal of risk factor research focusing on substance use and a wide variety of other negative outcomes is to inform the content and timing of efforts aimed at preventing the onset and/or escalation of these conditions (IOM, 1996; Mrazek & Haggerty, 1994). The added focus on informing the appropriate human *targets* of prevention has similarly emerged from decades of risk factor research demonstrating a wide variety of potent predictors (for a review see Hawkins *et al.*, 1992), but no single risk factor or constellation that is necessary or sufficient for the development of substance use, abuse, or dependence. Recent targeted prevention programs intended in full or in part to reduce rates of substance use have commonly aimed their efforts at "high-risk groups" defined by early signs/symptoms demonstrated in previous risk research [e.g. children of substance abusing parents (Kumpfer *et al.*, 1996), deviant youth (McMahon & Slough, 1996), and adolescents from impoverished socioeconomic backgrounds (Pizzolongo,

[1]Department of Psychology, Wesleyan University, Middletown, Connecticut.
[2]Mood and Anxiety Disorders Program, Intramural Research Program, Department of Health and Human Services, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland.
[3]University of Massachusetts, Amherst, Massachusetts.
[4]Division of Epidemiology, Services and Prevention Research, National Institute of Drug Abuse, Bethesda, Maryland.
[5]Correspondence should be directed to Dr. Lisa C. Dierker, Department of Psychology, Wesleyan University, 207 High Street, Middletown, Connecticut 06459; e-mail: ldierker@wesleyan.edu.

1389-4986/04/0900-0169/1 © 2004 Society for Prevention Research

1996)]. The fact that these programs often differ not only in terms of samples targeted, but also in basic program content and theoretical underpinnings, demonstrates the implicit expectation that different rather than common developmental pathways may explain the emergence of substance use behaviors among diverse population subgroups.

Extant risk research has most commonly inferred multiple pathways from a comparison of specific demographic groups defined by age, gender, or ethnicity. For example, deviance, other risk taking behaviors and poor school achievement have been found to be more strongly predictive of substance use behavior among middle school youth than among high school youth (Chassin *et al.*, 1991, 2000). Within studies examining gender differences in risk for substance use, susceptibility to the effects of parental smoking has been shown to be stronger among females (Curran *et al.*, 1999; Flay *et al.*, 1994; Kandel *et al.*, 1994; Robinson *et al.*, 1997), while deviance (Epstein *et al.*, 1999) and socioeconomic status (Curran *et al.*, 1999) have each been supported as stronger predictors of substance use among males. With regard to ethnicity, there has been rather consistent evidence for differences in the potency of interpersonal risk. For example, peer influence has been shown to be more important among Caucasian and Hispanic adolescents compared to African American adolescents (Flint *et al.*, 1998; Griesler & Kandel, 1998; Headen *et al.*, 1991; Landrine *et al.*, 1994; Maddahian *et al.*, 1988).

The relative importance of dividing the population into demographic subgroups (i.e. age, gender, and ethnicity) rather than groups defined by other factors that have been implicated as predictors of substance use is unclear. That is, the evaluation of demographic groups may or may not encompass the most meaningful splits for predicting the multiple etiologies at work in the population. Data-driven approaches that simultaneously examine a constellation of factors that show strong associations with substance use may yield more meaningful distinctions among individuals who use substances and in turn, begin to provide clues regarding the presence of multiple etiologic pathways.

There has also been recent consensus that the transition between stages of substance use represents a particularly important area within etiologic research that will more clearly allow us to define the specific risk factors at work throughout the development substance use behaviors (Glantz & Pickens, 1992; Jackson, 1997; Mayhew *et al.*, 2000; Prochaska & DiClemente, 1983). Research focused on stages of substance use have generally shown that the risk factors that predict initiation may differ from those that predict the maintenance or continuation of the behavior (Ary & Biglan, 1988; Glantz, 1992; Harrell *et al.*, 1998). Furthermore, recent work has suggested that initiation patterns may in fact differ for individuals who progress to substance dependence compared to those who do not (Mackesy-Amiti *et al.*, 1997).

The primary aim of the current study is to investigate whether individual risk factors or risk factor constellations can prospectively distinguish subgroups with elevated risk of substance use across various stages. Smoking outcomes were selected for the current investigation due to the fact that (1) cigarette smoking represents the substance use with one of the earliest onset ages (Kandel, 1982); (2) smoking is initiated before age 18 for the vast majority of individuals who will become nicotine dependent (Elders & Perry, 1994); and (3) as the leading cause of preventable deaths (CDC, 2002) smoking represents one of the most critical target outcomes for prevention. A diverse set of individual and contextual risk factors has been previously shown to be associated with smoking behavior, the most consistent of which include environmental exposure through peers (Gritz *et al.*, 1998; Jackson, 1997; Van Roosmalen & McDaniel, 1992), socioeconomic status (Chassin *et al.*, 1992), rebelliousness and antisocial behavior (Adalbjarnardottir & Rafnsson, 2002; Juon *et al.*, 2002), emotional distress (Allen *et al.*, 1994; Patton *et al.*, 1998; Pederson *et al.*, 1998), the accessibility of cigarettes (Forster *et al.*, 1997; Maddahian *et al.*, 1987), the use of other substances (Duncan *et al.*, 1998; Jackson *et al.*, 2002; Sussman *et al.*, 1997), self-esteem (Dielman *et al.*, 1988; Simon *et al.*, 1995), school achievement (Bryant *et al.*, 2001), and individual connectedness to the family and school environments (Conrad *et al.*, 1992; Resnick *et al.*, 1997). The present study set out to examine which factors might parsimoniously predict transition from nonsmoking to experimental smoking and regular smoking.

METHOD

Subjects

The National Longitudinal Study of Adolescent Health (Add Health) utilized cluster sampling to select and stratify adolescents by gender and grades 7 through 12 from 80 high schools and 52 feeder middle schools. These high schools were chosen from a

sample of all United States high schools that both included an 11th grade and enrolled 30 or more students. All students who completed the In-school Questionnaire plus those who did not complete a questionnaire but were listed on a school roster were eligible for selection into the In-home sample. Further, oversampling of four supplementary ethnic groups were drawn including African Americans, Chinese, Cuban, and Puerto Rican adolescents. In addition, the sample contains significant numbers of Mexican-Americans, Nicaraguans, Japanese, South Koreans, Filipinos, and Vietnamese. The study has been funded by the National Institute of Child Health and Human Development (NICHD) and 17 other federal agencies. Additional details about the survey design have been previously reported (Resnick *et al.*, 1997).

The present investigation is based on the 9,449 adolescents age 11–22 participating in both baseline and follow-up assessments, and who either maintained a nonsmoking or experimental smoking status or increased their level of smoking behavior across the 1-year assessment period. Data based on self-report information gathered during, in-home interviews were utilized in the present study. Interviewers employed by the National Opinion Research Center of the University of Chicago read and recorded responses for minimally intrusive or sensitive items while more sensitive information was gathered using computer-assisted interviewing (CASI).

The mean age of the selected sample at follow-up was 16.5, $SD = 1.6$ (99% of the sample was age 12–19), and 52% of the sample was female. In terms of race, 60% of the participants endorsed White, 25% African American, 3% American Indian, and 8% Asian. Seventeen percent of the sample endorsed Hispanic ethnicity.

Measures

Levels of smoking were defined for each subject at both the baseline (lifetime) and follow-up (1 year) assessment based on the questions "Have you (ever/since month of last interview) tried cigarette smoking, even just one or two puffs?" and "Have you (ever/since month of last interview) smoked cigarettes regularly, that is at least one cigarette every day for 30 days?" At each assessment, subjects were categorized as *nonsmokers* if they responded negatively to having tried cigarette smoking, as *experimenters* if they responded positively to having tried cigarette smoking, but negatively to daily smoking for 30 days, and as *regular smokers* if they responded positively to both having tried cigarette smoking and to daily smoking for 30 days.

Groups were created to reflect prospective increases in smoking behavior. Comparisons included (1) nonsmokers at baseline who were experimenters by the follow-up assessment ("experimenters") vs. nonsmokers at baseline who remained nonsmokers at follow-up ("nonsmokers"); (2) nonsmokers at baseline who were regular smokers by the follow-up assessment ("rapid regular smokers") vs. nonsmokers at baseline who were experimenters by follow-up ("experimenters"); and (3) experimenters at baseline who became regular smokers by the follow-up assessment ("regular smokers") vs. experimenters at baseline who remained experimenters at follow-up ("continued experimenters").

Variables represented in Add Health that have been previously established as correlates of smoking behaviors were included as predictor variables for the present investigation. A total of 26 predictor variables from the baseline assessment were selected. Youth characteristics included age and gender. Race was also evaluated through separate questions and the potential for the endorsement of multiple categories including White, African American, Asian, American Indian and a single question assessing Hispanic ethnicity. Poverty status for each adolescent was coded as a dichotomous variable based on whether the adolescent's mother or father was currently on public assistance.

Baseline alcohol use was assessed with the question, "Did you ever drink beer, wine or liquor when you were not with your parents or other adults in your family?" Similar dichotomous variables were created for the presence or absence of marijuana and cocaine use. Variables representing alcohol problems in the past 12 months were created as a frequency count based on dichotomizing to the presence or absence of problems that adolescents indicated were associated with drinking (i.e. trouble with parents, problems with school or schoolwork, problems with friends, problems with someone the adolescent is dating, regretting participation in a sexual situation and getting into one or more physical fights). For adolescents who at baseline had smoked at least one cigarette, variables indicating smoking quantity and frequency over the past 30 days were also created. Frequency indicated the number of days smoked during the past 30 days and quantity indicated the number of cigarettes smoked per day over the same time period.

Emotional and behavioral indices included continuous measures developed by Resnick and

colleagues (Resnick *et al.*, 1997). Deviance included 11 items assessing the presence and frequency (0 = *never*, 1 = *1* or *2 times*, 2 = *3* or *4 times*, and 3 = *5* or *more times*) of past year behaviors regarding vandalism, other property damage, lying, stealing, running away, driving without permission, selling drugs, and skipping school. Employing an identical rating scale, violence included eight items assessing physical fighting, pulling a knife or gun on another individual, shooting or stabbing someone, hurting someone badly enough to require bandages or professional care, threatening or using a weapon to get something from someone, group violence, and the use of a weapon in a fight during the past year. Further, a depression scale included 17 items assessing past week symptoms of negative emotionality and a self-esteem scale included 10 items addressing a positive personal and physical self-image (Resnick *et al.*, 1997). Similar continuous measures of family environment included parental presence (frequency with which a parent is home when the adolescent leaves for school, returns from school and goes to bed), parent–child activities (number of activities with parents in the past 4 weeks), and family connectedness (13 items assessing the adolescent's relationship with and feelings toward parents and family members; Resnick *et al.*, 1997). An additional dichotomous variable indicating the accessibility of cigarettes was also included based on the question "Are cigarettes easily available to you in your home?"

School constructs included expulsion ("Have you ever been expelled from school?") school connectedness (8 items assessing the adolescent's relationship with teachers and peers and perceptions of school atmosphere) and an approximate Grade Point Average calculated on a 4.0 scale according to the adolescent's reported performance in up to four school subjects. Finally, peer smoking was included as a count from 0 to 3 based on the question "Of your 3 best friends, how many smoke more than one cigarette per day?"

Statistical Approach

First, bivariate analyses were conducted for each predictor variable with regard to the three paired outcome comparisons defined above (chi-square and ANOVA for categorical and continuous variables, respectively). Next, Classification and Regression Tree (CART) analyses were performed to test linear and nonlinear relationships among the predictor variables. CART represents a nonparametric analytic technique that makes binary splits in a sample that separates the subjects into the purest high-risk and low-risk subgroups with respect to the outcome. All possible separations (categorical) or cutpoints (continuous) are tested by comparing the chi-square statistic in relation to the outcome (Breiman *et al.*, 1984). If a variable such as age (i.e. a continuous variable) is considered, 12 distinct ages ranging from 11 to 22 years were recorded in the data set, and therefore, the entire sample may be split in 11 different ways. Each of these 11 possible splits is then tested according to which separates the subjects into the purest possible high-risk and low-risk subgroups (Zhang & Bracken, 1996). CART generates a decision tree in which each of the variables in the model is assessed at each decision point including the variables on which the previous splits were made. This analytic tool was selected because it is particularly suited to an examination of subgroups in that the models potentially generate multiple discrete groups with high or low proportions of individuals with the outcome of interest. For the present analyses, various available "goodness of split" criteria were used including Gini, Systematic Gini, Twoing, and Power Modified Twoing.

Following the initial growing of the tree, a cross-validation procedure was used to protect against an overfit model by using a random subset of the data and retaining "branches" of the tree that improve the correct classification rate (Breiman *et al.*, 1984). Briefly, CART divides the data into learning and test subsamples. The learning sample is used to grow an overly large tree, while the test sample is then used to estimate the rate at which cases are misclassified. The misclassification rate is calculated for every sized tree and the selected subtree represents the lowest probability of misclassification (Steinberg & Colla, 1995). Following the creation of the classification tree, individual sensitivities (true positive rates) and specificities (true negative rates) were calculated for the overall model and for individual branches to determine the contribution of each risk factor or in the case of complex branching, constellation of factors, in accurately classifying cases. Classification tree models were further investigated through traditional logistic regression analyses.

Because of the fact that CART attempts to maximize sensitivity and specificity with the simplest tree structure, it is possible for variables that do not represent primary splits in the model to be of notable importance in the prediction of outcome. When variables are for example highly correlated or provide similar information (e.g. mother's education and father's

education; Steinberg & Colla, 1995), only one is likely to be selected for the model. The absence of the alternate variable from the model does not then suggest that it is unimportant, but rather that it is "masked" by the other. To handle this phenomenon of masking, an importance measure was also calculated for each variable that was not selected as a primary predictor. The importance measure reflects the improvement that would be attributable to each variable in its role as a surrogate to the primary split. The values of these improvements are summed over each node and totaled, and are scaled relative to the best performing variable (Steinberg & Colla, 1995). The variable with the highest sum of improvements is scored 100, and all other variables will have lower scores ranging downward to zero. In short, the importance score measures a variable's ability to mimic the chosen tree and to play the role as stand-in for variables appearing as primary splits. To evaluate masking in the present models, variables with importance scores ranging from 15 to 100 are presented.

RESULTS

Of those adolescents who were nonsmokers at baseline ($n = 5,401$), 1,068 (17%) had initiated smoking prior to the second survey. Of those who had initiated by follow-up, 847 (79%) remained at the experimental level, while the remaining 221 (21%) rapidly progressed to regular smoking. Among baseline experimenters who continued smoking during the period proceeding the follow-up survey ($n = 2,980$), 2,148 (72%) remained experimental smokers, while 832 (28%) progressed to regular smoking. Table 1 presents the bivariate associations between baseline characteristics and change in smoking status. Bonferoni adjustment (Westfall *et al.*, 1999) of the p value for the evaluation of these multiple comparisons is $p < .002$.

Nonsmoking to Experimentation

Chi-square analyses for categorical variables and Analysis of Variance for continuous variables yielded the following differences: First, compared with adolescents who remained nonsmokers at follow-up, those who progressed to experimentation were older, $F(1, 6246) = 12.9$, $p < .0003$, and exhibited more substance use and related symptoms including alcohol use without supervision ($\chi^2 = 60.0$, $df = 1$, $p < .0001$), marijuana use ($\chi^2 = 43.8$, $df = 1$, $p < .0001$), and one or more problems associated with alcohol use ($\chi^2 = 28.9$, $df = 1$, $p < .0001$). Further, experimenters reported more violence, $F(1, 6238) = 40.6$, $p < .0001$, deviance, $F(1, 6238) = 71.6$, $p < .0001$, and depression, $F(1, 6243) = 11.2$, $p < .0008$, than their nonsmoking counterparts and were more likely to have cigarettes easily available in their home ($\chi^2 = 12.1$, $df = 1$, $p < .0005$). In relation to school, experimenters were found to have more smoking friends, $F(1, 6147) = 40.1$, $p < .0001$, a lower grade point average, $F(1, 6106) = 16.7$, $p < .0001$, and less connectedness to school, $F(1, 6183) = 26.3$, $p < .0001$; Table 1.

Next, baseline variables were included as possible contributors to a CART model evaluating smoking experimentation at follow-up (Fig. 1). Each pentagon represents a decision point. For each decision point, the predictor variable and cut point are presented. Final groups with high and low outcome probability are represented by rectangles in the figure and include outcome frequencies and percentages. Shaded rectangles represent subgroups with relatively high rates of smoking, nonshaded rectangles, relatively low rates.

The deviance score was the first variable to separate the sample into two subgroups. Adolescents with a baseline deviance score greater than 0.112 (Sample range 0–2.8; $M = 0.13$, $SD = 0.209$) were more likely to have progressed to smoking experimentation compared to adolescents not meeting this cutoff (18.6% vs. 11.2%). Because the deviance score represents the average of behavioral frequency across the 11 deviance items, for descriptive purposes, the average number of symptoms endorsed by this group was also examined. Adolescents with a deviance score of greater than 0.112 reported an average of 2.5 ($SD = 1.43$) deviant behaviors, with 80% of the sample reporting two or more, and only 18% reporting more than three.

Of the adolescents with deviance scores less than or equal to 0.112, a further subdivision was made with the dichotomous variable of alcohol use without supervision. Adolescents who reported having used alcohol without supervision at baseline were more likely to have proceeded to smoking experimentation at follow-up compared to those who did not (21.1% vs. 10.3%). The total CART model classified 63% of the sample correctly, 52% of experimenters (sensitivity) and 65% of nonsmokers (specificity). Sensitivity calculated for each branch separately revealed that the deviance branch identified 44% of experimenters, while the alcohol branch identified the additional 8%. The evaluation of variable importance for predictors not included in the present model revealed marijuana

Table 1. Baseline Characteristics by Longitudinal Smoking Levels[a]

	Nonsmoker					Nonsmoker					Experimenter				
	Non ($N = 5,401$)		Exp ($N = 847$)			Exp ($N = 847$)		Reg ($N = 221$)			Exp ($N = 2,148$)		Reg ($N = 832$)		
Baseline characteristics	*N/M*	*%/SD*	*N/M*	*%/SD*	*p* value	*N/M*	*%/SD*	*N/M*	*%/SD*	*p* value	*N/M*	*%/SD*	*N/M*	*%/SD*	*p* value
Gender (% female)	2,838	52.6%	427	50.4%	.25	427	50.4%	80	36.2%	.0002	1,101	51.3%	443	53.3%	.33
Age: M (SD)	*15.5*	*1.66*	*15.3*	*1.66*	*.0003*	*15.3*	*1.66*	*15.6*	*1.61*	*.02*	*15.6*	*1.56*	*15.7*	*1.52*	*.21*
Poverty (%)	572	10.6%	103	12.2%	.17	103	12.2%	30	13.6%	.57	223	10.4%	94	11.30%	.47
Race/Ethnicity															
White (%)	2,998	55.6%	516	61.1%	.003	516	61.1%	158	71.5%	.0041	1,404	65.6%	606	72.9%	.0001
African American (%)	1,553	28.8%	204	24.1%	.0052	204	24.1%	39	17.7%	.04	421	19.7%	113	13.6%	.0001
American Indian (%)	153	2.8%	29	3.4%	.34	29	3.4%	9	4.1%	.65	93	4.4%	36	4.4%	.99
Asian (%)	494	9.2%	63	7.5%	.1062	63	7.5%	9	4.1%	.047	137	6.4%	52	6.3%	.89
Hispanic Ethnicity (%)	897	16.7%	147	17.4%	.59	147	17.4%	35	15.9%	.60	381	17.8%	133	16.0%	.24
Substance use															
Smoking frequency	—	—	—	—	—	—	—	—	—	—	1.9	3.98	4.8	*7.33*	.0001
Smoking quantity	—	—	—	—	—	—	—	—	—	—	0.5	1.90	1.5	*3.76*	.0001
Alcohol use w/o permission	711	13.2%	197	23.3%	.0001	197	23.3%	70	31.7%	.01	1,165	54.3%	501	60.2%	.0036
Alcohol problems (≥1) (%)	238	4.4%	74	8.7%	.0001	74	8.7%	33	14.9%	.0063	537	25.0%	273	32.8%	.0001
Marijuana use (%)	306	5.7%	99	11.7%	.0001	99	11.7%	54	24.4%	.0001	712	33.2%	345	41.5%	.0001
Cocaine use (%)	16	0.3%	2	0.2%	.76	2	0.2%	4	1.9%	.0049	12	0.6%	9	1.1%	.12
Emotional and behavioral															
Violence	*0.1*	*0.23*	*0.2*	*0.31*	*.0001*	*0.2*	*0.31*	*0.3*	*0.38*	*.0010*	*0.2*	*0.32*	*0.3*	*0.41*	*.0001*
Deviance	*0.1*	*0.20*	*0.2*	*0.25*	*.0001*	*0.2*	*0.25*	*0.3*	*0.38*	*.0001*	*0.3*	*0.31*	*0.4*	*0.37*	*.0001*
Depression	*0.4*	*0.35*	*0.5*	*0.36*	*.0008*	*0.5*	*0.36*	*0.5*	*0.40*	*.2960*	*0.5*	*0.39*	*0.6*	*0.46*	*.0001*
Self-esteem	*4.2*	*0.51*	*4.1*	*0.53*	*.0080*	*4.1*	*0.53*	*4.1*	*0.55*	*.5195*	*4.1*	*5.30*	*4.0*	*0.59*	*.0001*
Family															
Cigarettes available (%)	1,097	20.4%	216	25.6%	.0005	216	25.6%	77	35.2%	.0049	672	31.3%	283	34.1%	.15
Family connectedness	*22.9*	*2.71*	*22.7*	*2.57*	*.14*	*22.7*	*2.57*	*21.9*	*3.60*	*.0002*	*22.3*	*2.83*	21.8	*3.19*	*.0001*
Parent activities	*6.5*	*3.41*	*6.4*	*3.44*	*.49*	*6.4*	*3.44*	*5.8*	*3.38*	*.0136*	*6.3*	*3.35*	*6.0*	*3.41*	*.03*
Parental presence	*13.6*	*2.01*	*13.5*	*2.00*	*.43*	*13.5*	*2.00*	*13.2*	*2.46*	*.0300*	*13.3*	*2.16*	*13.4*	*2.06*	*.37*
School															
History of expulsion (≥1)	119	2.2%	38	4.5%	.0001	41	4.6%	13	5.9%	.45	78	3.6%	37	4.5%	.30
GPA	*3.0*	*0.74*	*2.9*	*0.73*	*.0001*	*2.9*	*0.73*	*2.6*	*0.74*	*.0001*	*2.8*	*0.75*	*2.6*	*0.72*	*.0001*
School connectedness	*3.7*	*0.58*	*3.6*	*0.60*	*.0001*	*3.6*	*0.60*	*3.4*	*0.73*	*.0001*	*3.5*	*0.60*	*3.4*	*0.64*	*.0001*
Friends smoking	*0.3*	*0.68*	*0.5*	*0.81*	*.0001*	*0.5*	*0.81*	*1.1*	*1.19*	*.0001*	*0.7*	*0.93*	*1.2*	*1.06*	*.0001*

[a] Rows in *italics* represent comparisons based on Analysis of Variance. Means and standard deviations are presented. Rows in nonitalics represent comparisons based on chi-square analyses ($df = 1$). Frequencies and percentages are presented. Bonferroni adjustment for each of the three sets of comparisons is $p < .002$.

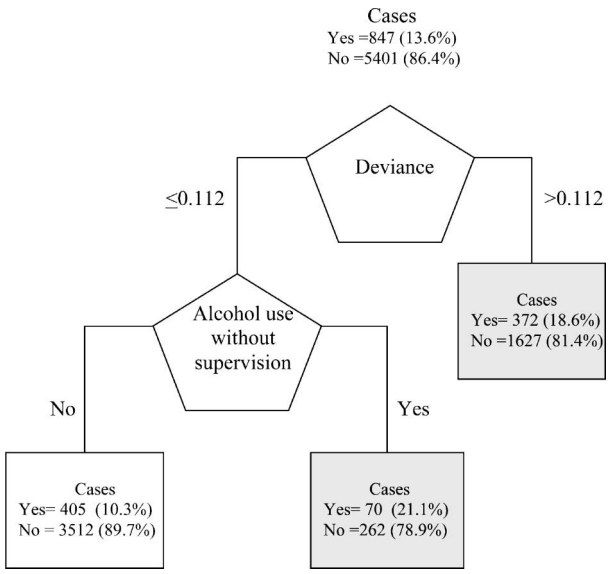

Fig. 1. Classification tree ($N = 6{,}248$) for progression to experimental smoking at follow-up.

use (Importance score = 43.9) and problems associated with alcohol use (Importance score = 32.9) to be the closest competitors. Neither, however, substantially increased the predictive value of the model when variables represented primary splits were removed from the model. Further, deviance and alcohol use without permission remained primary splits when various "goodness of split" criteria were implemented.

Next, logistic regression analyses were used to examine differences between cases (true positives) from each branch of the model (i.e. the deviance type smoking experimenters vs. the alcohol type smoking experimenters). All predictor variables with the exception of alcohol use without supervision and deviance scores were included in the model (Table 2). Deviant type smoking experimenters were found to differ from the alcohol type smoking experimenters in that the deviant group exhibited more symptoms of depression [OR = 3.2: CI (1.06–9.45), unadjusted: $M = 0.6$, $SD = 0.37$ vs. $M = 0.4$, $SD = 0.34$], had a lower GPA [OR = 0.4: CI (0.27–0.72), unadjusted: $M = 2.8$, $SD = 0.75$ vs. $M = 3.1$, $SD = 0.67$] and were less likely to be African American [OR = 0.1: CI (0.04–0.59), unadjusted rates: 22% vs. 31%].

Nonsmoking to Regular Smoking

The next set of analyses focused on differences between nonsmokers at baseline who progressed rapidly to regular smoking at follow-up versus those who increased their smoking only to experimental levels. On the basis of chi-square and Analysis of Variance, rapid escalators were more often male ($\chi^2 = 14.2$, $df = 1$, $p < .0002$), more likely at baseline to have used marijuana ($\chi^2 = 43.8$, $df = 1$, $p < .0001$), exhibited more violence, $F(1, 1066) = 10.6$, $p < .001$, and deviance, $F(1, 1065) = 20.5$, $p < .001$, shown less connectedness to family, $F(1, 1066) = 13.7$, $p < .0002$, and school, $F(1, 1050) = 17.4$, $p < .0001$, reported a lower GPA, $F(1, 1043) = 25.5$, $p < .0001$, and more smoking friends, $F(1, 1044) = 71.5$, $p < .0001$, compared to experimenters (Table 1).

The CART model examining rapid escalation to regular smoking split the sample on friends' smoking for the initial branching (Fig. 2). Lower rates of regular smoking were observed among adolescents with two or fewer smoking friends at baseline compared to adolescents reporting that all three of their best friends smoked more than one cigarette per day (17.1% vs. 57.5%). Among adolescents with two or fewer smoking friends, a second split in the model was made. Higher rates of regular smoking were seen among adolescents with a grade point average less than or equal to 2.7 compared to those with a higher grade point average (24.1% vs. 13.3%). The overall model classified 61% of the sample correctly, 65% of regular smokers (sensitivity) and 59% of experimenters. Sensitivity calculated for each branch separately revealed that the friends' smoking branch identified 21% of regular smokers, while the GPA branch identified 44%. The further evaluation of variable importance for predictors not included in this model revealed only school connectedness as a distant competitor (Importance score = 23.6). Again, friend smoking and grade point average remained primary splits when various "goodness of split" criteria were implemented.

Multivariate logistic regression analyses used to examine differences between cases (true positives) from each branch of the model revealed that smoking friend type rapid escalators were more likely at baseline to have used marijuana [OR = 5.7: CI (1.63–20.12), unadjusted rates: 52.2% vs. 15.8%] and to have had one or more problems associated with alcohol use [OR = 5.8: CI (1.48–22.55), unadjusted rates: 39.1% vs. 9.9%; Table 2].

Experimentation to Regular Smoking

When attempting to predict regular smoking among adolescents who were experimenters at

Table 2. Logistic Regression Analyses Comparing True Positive Cases Within Each Branch of the Classification Models

	Deviant type experimenters ($n = 372$) vs. Alcohol type experimenters ($n = 70$)		Friend type rapid escalators ($n = 46$) vs. GPA type rapid escalators ($n = 98$)		Frequency type regular smokers ($n = 227$) vs. Friend type regular smokers ($n = 374$)	
	OR	CI	OR	CI	OR	CI
Gender (% female)	0.9	0.45–1.66	0.8	0.26–2.53	0.8	0.46–1.23
Age: M (SD)	0.9	0.75–1.12	0.9	0.64–1.23	1.0	0.81–1.11
Poverty (%)	1.6	0.46–5.29	0.5	0.13–1.71	1.19	0.56–2.52
Race/Ethnicity						
White (%)	0.4	0.11–1.39	2.8	0.5–16.04	2.7*	1.09–6.69
African American (%)	0.1**	0.04–0.59	0.7	0.08–5.35	1.7	0.57–4.78
American Indian (%)	0.8	0.16–4.54	0.7	0.10–4.86	0.6	0.15–2.07
Asian (%)	0.8	0.18–3.16	9.4	0.41–215.7	2.2	0.71–6.82
Hispanic Ethnicity (%)	1.4	0.53–3.93	1.16	0.31–4.43	0.8	0.4–1.66
Substance use						
Smoking frequency	—	—	—	—	2.6	2.12–3.20
Alcohol use w/o permission	—	—	1.0	0.28–3.34	0.7	0.39–1.15
Alcohol problems (≥1) (%)	1.2	0.49–2.74	5.8**	1.48–22.55	1.3	0.87–2.22
Marijuana use (%)	0.6	0.25–1.34	5.73**	1.63–20.12	2.2*	1.35–3.70
Cocaine use (%)	0.3	0.02–4.52	0.5	0.06–4.92	0.3	0.09–1.20
Emotional and behavioral						
Violence	3.5	0.90–13.28	3.0	0.64–14.28	0.4*	0.17–0.75
Deviance	—	—	0.4	0.08–2.22	3.7*	1.70–8.21
Depression	3.2*	1.06–9.45	2.9	0.61–13.59	1.0	0.54–1.82
Self-esteem	0.9	0.43–1.81	1.7	0.63–4.90	1.0	0.61–1.64
Family						
Cigarettes available (%)	1.1	0.52–2.17	1.9	0.73–5.21	0.8	0.49–1.28
Family connectedness	0.9	0.75–1.06	1.0	0.83–1.18	1.0	0.88–1.08
Parent activities	1.0	0.89–1.07	1.0	0.87–1.18	1.0	0.93–1.07
Parental presence	1.0	0.87–1.17	1.0	0.82–1.12	1.1	0.98–1.27
School						
History of expulsion (≥1)	4.3	0.47–39.60	1.4	0.25–7.62	1.7	0.57–5.17
GPA	0.4**	0.27–0.72	—	—	1.2	0.80–1.65
School connectedness	0.9	0.50–1.71	1.0	0.50–2.26	1.1	0.70–1.67
Friends smoking	0.8	0.57–1.19	—	—	—	—

*$p < .05$, **$p < .01$.

baseline versus those who continued smoking at an experimental level, quantity and frequency of baseline smoking were also evaluated, based on previous research that has implicated past smoking behaviors as the most potent predictor of future smoking behaviors. The comparison between regular and experimental smokers revealed a total of 14 significant bivariate associations (Table 1). Regular smokers at follow-up were more likely to be White ($\chi^2 = 14.9$, $df = 1$, $p < .0001$), less likely to be African American ($\chi^2 = 14.9$, $df = 1$, $p < .0002$), and to have smoked both more, $F(1, 2952) = 83.0$, $p < .0001$, and more often, $F(1, 2150) = 142.4$, $p < .0001$, in the past month at the baseline assessment than experimental smokers. Regular smokers were also more likely to report one or more associated problems with alcohol use ($\chi^2 = 18.5$, $df = 1$, $p < .0001$). Furthermore, they were more likely to use marijuana ($\chi^2 = 18.1$, $df = 1$, $p < .0001$), as well as to exhibit more violence, $F(1, 2974) = 24.2$, $p < .0001$, deviance, $F(1, 2977) = 28.8$, $p < .0001$, and depression, $F(1, 2977) = 27.0$, $p < .0001$, then experimental smokers. Finally, regular smokers were found to report lower self-esteem $F(1, 2977) = 24.8$, $p < .0001$, lower school, $F(1, 2934) = 28.3$, $p < .0001$, and family connectedness, $F(1, 2978) = 21.2$, $p < .0001$), a lower GPA, $F(1, 2919) = 32.9$, $p < .0001$, as well as more smoking friends, $F(1, 2943) = 142.2$, $p < .0001$, than their experimental counterparts.

Notably, the final CART analysis revealed the same number of subgroups as previous models (Fig. 3). However, smoking frequency (>4 1/2 days smoking in past 30) was the first split followed by friend's smoking (1 to 3 smoking friends). Adolescents reporting baseline smoking at least 4 ½ days out of the past 30 were more likely to have progressed

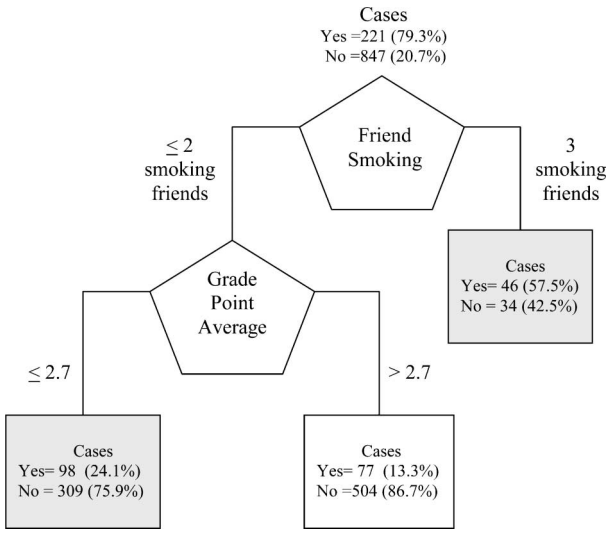

Fig. 2. Classification tree ($N = 1{,}068$) for progression to experimental smoking at follow-up.

to regular smoking compared to adolescents smoking fewer than 4 1/2 days (53.5% vs. 23.4%). For the adolescents smoking *fewer* than 4 1/2 days in the past 30, having one or more smoking friends at baseline increased their risk of becoming regular smokers at follow-up compared to those with no smoking friends (32.0% vs. 16.3%). The model classified 59% of the sample correctly, 72% of regular smokers (sensitivity) and 54% of experimenters (specificity). Sensitivity calculated for each branch separately revealed that the smoking frequency branch identified 27% of regular smokers, while the friend's smoking branch identified 45%. The evaluation of variable importance for predictors not included in this final model revealed smoking quantity (Importance score = 78.7) and school connectedness (Importance score = 15.8) to be the closest competitors. When smoking frequency was removed from the list of predictors, the primary split in the model became smoking friends, and the resulting model, while showing comparable sensitivity and specificity was also much more complex than that including smoking frequency. Again, smoking frequency and friends' smoking remained primary splits when various "goodness of split" criteria were tested.

Logistic regression analyses used to examine differences between true cases from each branch showed that high frequency type regular smokers smoked at baseline in higher quantities [OR = 2.6: CI (2.12–3.20), unadjusted: $M = 4.1$ $SD = 6.30$ vs. $M = 0.6$ $SD = 1.27$], were less violent [OR = 0.4: CI (0.17–0.75), not demonstrated in unadjusted means] and more deviant [OR = 3.7: CI (1.70–8.21), unadjusted: $M = 0.5$ $SD = 0.42$ vs. $M = 0.3$ $SD = 0.36$] than the smoking friend type regular smokers. Further, the high frequency type was more likely to be White [OR = 2.7: CI (1.09–6.69), unadjusted rates: 77.4% vs. 71.9%] and more likely to have smoked marijuana [OR = 2.2: CI (1.35–3.70), unadjusted rates: 59.5% vs. 35.8%].

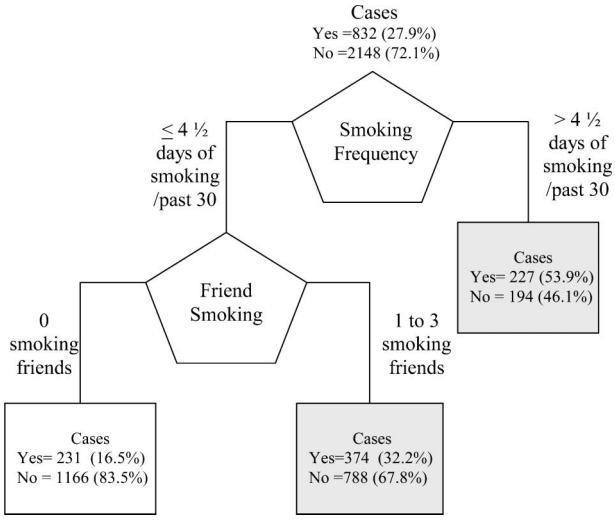

Fig. 3. Classification tree ($N = 2{,}980$) for progression to experimental smoking at follow-up.

DISCUSSION

This study employed longitudinal data from a large sample of adolescents to investigate the presence of population subgroups with increased risk for progression to different smoking stages. Overall, results indicated general success in sensitively identifying escalation to experimental and regular smoking over the 1-year assessment period. For each level of smoking, a relatively simple model including two subgroups predicted over half of the smoking cases and these two-group models identified higher rates of regular smokers compared to experimental smokers.

Nonsmoking to Experimentation

Over half of the adolescents who proceeded to smoking experimentation by follow-up were correctly classified based on the creation of two risk groups defined by deviant behavior and unsupervised

alcohol use. The presence of deviance, most often including the endorsement of only two or three related behaviors, identified a remarkable 44% of smoking experimenters. This finding strongly suggests that even relatively low levels of deviant behavior confer an important increase in risk, thus adding to abundant evidence suggesting the prominent role of deviant behavior in the etiology of smoking (Brown *et al.*, 1996; Griesler & Kandel, 1998; Sieber & Angst, 1990) and supporting a general construct of problem behaviors, a theory extensively discussed in previous literature (Jessor & Jessor, 1977). Although deviant behavior has commonly been used to target youth for intervention, predominantly, these programs have focused on younger children as well as those exhibiting higher thresholds of deviant behavior. (Conduct Problems Prevention Research Group, 2000; The Fast Track Program, 1992). The present findings suggest that adolescents engaging in these relatively low levels of deviant behavior (i.e. levels that classify 1/3 of the population of baseline nonsmokers) represent an important target for intervention efforts aimed at preventing smoking onset and escalation among nonsmokers.

The addition of the alcohol type risk group for smoking experimentation (i.e. unsupervised alcohol use and very little or no deviant behavior) correctly classified a further 8% of smoking experimenters. The identification of an alcohol use group, separate from those exhibiting the well-articulated deviant behavior syndrome suggests the presence of a more behaviorally normative group of adolescents at substantial risk for smoking experimentation. Aside from exhibiting few or no other externalizing symptoms, alcohol type smokers were also found to enjoy generally higher functioning, based on better school performance and fewer symptoms of depression, compared to the deviant type smokers. It is possible that this group will be less likely to continue in their escalation of smoking behaviors and would not be represented in subgroup analyses focusing on higher levels of use.

Indeed, the difference in depression scores between deviant type smokers and alcohol type smokers did not provide direct support for the "dual pathway" to substance use hypothesis through depression and other internalizing symptoms on the one hand and deviant behavior on the other (Cloninger *et al.*, 1981, 1988; Rohde *et al.*, 1991, 1996; Zucker, 1994). Rather, while the present study may confirm that depression confers an increased risk of smoking behavior mainly in the presence of comorbid deviance (Miller-Johnson *et al.*, 1998), the need for more direct empirical attention to the dual pathways hypothesis is also implicated. That is, complications regarding the characterization of depression (i.e. as a diagnostic entity or continuum of symptoms), and the lack of assessment periods that encompass the entire age of risk for initiation and progression limit our ability to provide adequate evidence on this topic. For example, there is retrospective evidence to suggest that the depression pathway is one that is largely followed in later onset substance users (Cloninger *et al.*, 1988; Zucker, 1994). Its emergence in pattern-centered analyses may require the inclusion of prospective measurement into young adulthood.

Finally, the alcohol type smoking experimenters were also more likely to be African American, compared to their deviant type counterparts. In recent years, the possibility that different racial or ethnic groups may travel along distinct pathways to and through smoking stages has been given substantial and focused attention (Alexander *et al.*, 1999; Botvin *et al.*, 1992; Headen *et al.*, 1991; Robinson *et al.*, 1997). This issue has been of particular interest for smoking researchers based on the consistently higher rates of smoking among White compared to Black adolescents (Feigelman & Lee, 1995), in contrast to higher rates of smoking related illness among Black adults (USDHHS, 1998). However, differences in rates of African American youth between the deviant and alcohol types smokers, while significant, were not substantial (22% vs. 31%). Furthermore, previous research has clearly demonstrated deviant behavior to be a prominent risk factor for smoking among both Black and White adolescents (USDHHS, 1994). Clearly, racial and ethnic factors will remain a critical consideration in future research.

Nonsmoking to Regular Smoking

In an attempt to predict relatively rapid escalation to regular smoking, analyses were conducted comparing regular smokers to experimental smokers at follow-up who had been nonsmokers at baseline. The largest proportion of rapid regular smokers were captured by the criterion of a relatively low grade point average (i.e. lower than B) among youth with two or fewer smoking friends at baseline. Again, this single branch of the model identified a striking 44% of regular smokers. Further, having three or more smoking friends at baseline identified an additional 21% of rapid regular smokers. Taken together, these results confirm previous findings implicating exposure to other smokers and reported school performance as

two of the strongest predictors of established smoking patterns (Boyle *et al.*, 1993; Choi *et al.*, 1997) and also expand on this work by defining separate subgroups at high risk for rapid escalation to regular smoking.

The further consideration of factors that distinguish between rapid regular smokers within these two subgroups demonstrated that regular smokers who had at least three smoking friends at baseline were also more likely to have used marijuana and to exhibit one or more problems associated with alcohol use compared to the regular smokers defined by relatively poor school performance. Thus, it is likely that the smoking friends type regular smoker experiences elevated risk for smoking through shared beliefs, values, and behaviors regarding multiple forms of drug use within their peer cluster, as described by one of the leading theories in the drug use literature (Oetting & Beauvais, 1987). Notably, the measurement of peer smoking included a respondent-report assessment of the number of friends engaging in *regular* smoking rather than experimental use.

Experimentation to Regular Smoking

Nearly three quarters of adolescents who progressed from experimental smoking to regular smoking during the study's 1-year assessment period were identified based on the creation of two population subgroups: (1) Smoking more than 4 ½ days during the 30 days prior to baseline and (2) Smoking 4 ½ days or fewer over those 30 days, and reporting at least one smoking friend. Again, almost half of the regular smokers were accurately classified by the smoking friends' branch of the model, showing that even infrequent experimentation in the context of peer smoking puts youth at increased risk for established smoking patterns (Pierce *et al.*, 1996). Clearly, these potentially "weekend only" smokers remain important targets for prevention and may again, represent youth that escalate to regular smoking through forces inherent in the peer cluster (Oetting & Beauvais, 1986).

Although the higher frequency smoking group identified only 27% of regular smokers, it also brought with it a substantially lower false positive rate, in that over half (58%) of the higher frequency smokers at baseline had become regular smokers within 1 year. This group seems most consistent with previous work showing that established experimenters have the highest risk for progressing to regular smoking (Choi *et al.*, 1997). In that the high frequency type smoker could also be distinguished from the smoking friend type through higher quantities of smoking at baseline, more deviance, and a higher likelihood of marijuana use, this group likely represents youth who will require intervention that focuses on a range of problem behaviors. More extensive longitudinal research is needed to determine whether these high frequency type smokers best represent those youth whose experimental smoking stage can be characterized by deviant behavior, thus adding to our ability to hypothesize smoking pathways.

Strengths and Limitations

The strengths of Add Health for addressing the present research questions are numerous. First, Add Health included an extremely large sample embodying the most informative age group for the mapping of change in smoking progression. Second, Add Health included the assessment of an extremely wide array of individual and contextual factors believed to play a role in smoking etiology. Third, the longitudinal design and measurement of smoking stages based on quantity and frequency of use allowed for an evaluation of factors associated with movement to separate stages of smoking behavior. Similar to previous work, the present findings suggested that in fact, primary risk factors may differ for predicting smoking initiation, maintenance, and movement to established levels of use (Ary & Biglan, 1988; Harrell *et al.*, 1998). Finally, various racial/ethnic groups were oversampled (as well as disabled adolescents and sibling pairs) allowing for increased power to detect the possibility of important differences based on ethnicity. Because a nationally representative sample may mask characteristics of poorly represented groups, the choice to use the unweighted sample improved our ability to address the present research questions. It must be noted however, that this strength also brings with it an important caution about the generalizability of the present results. That is, while subgroups have been identified within a large group of adolescents sampled across the United States, the clustered nature of the sampling approach and the oversampling of special groups of interest also mean that the results should not be construed as nationally representative estimates.

This study also drew its strength from the use of a pattern oriented analytic procedure to discover factors associated with smoking risk (Bates, 2000). Classification and regression tree modeling was selected due to its potential to fit numerous interactions that

cannot be handled as efficiently in regression analyses, as well as its data-driven nature creating decision points yielding groups with low or high probability on the outcome of interest. In that CART modeling assumes that the nature of variables are basically binary, it is possible that further complexities are present in the data that are not explored by the binary structure. Thus, while CART remains a useful approach for investigating the presence of subgroups within the population, other pattern-centered approaches should also be considered in future research.

The present study did evaluate a wide array of potential predictors associated with smoking, yet, as is always the case in risk research, additional factors that have been identified in the literature were not examined and should be included in future research efforts. For example, parental smoking was excluded from the analyses due to the incomplete measurement of this construct. That is, while the parent participating in the home-based interview was asked a single question about smoking behavior, the variety of potential respondents as well as the marked lack of father report made this variable comparably weak. This methodologic issue, coupled with accumulating evidence that parental smoking has only small or nonsignificant effects on offspring smoking (Conrad *et al.*, 1992; Tyas & Pederson, 1998), particularly when more influential factors like peer smoking are included in analytic models (Ary & Biglan, 1988; Mittelmark *et al.*, 1987), supports our exclusion of this variable in the current paper.

Although a large survey such as Add Health presents an opportunity to examine diverse constructs, the measurement of those constructs is also commonly limited to brief scales or single-item indicators that do not necessarily represent the most reliable and valid options. This should be considered when interpreting the present findings. Further, it should be noted that statistically significant findings at the bivariate level were often small and not necessarily suggestive of clinically meaningful differences. Overall patterns were consistent with previous literature and are provided as descriptive information for the evaluation of the classification models. Finally, the decision to analyze transitions related to smoking progression necessitated using a portion of the Add Health data. Thus, related to progression, adolescents who were already regular smokers at baseline were not evaluated for their smoking risk. This subsample may represent a group of very early starters or simply a cohort (i.e. regular prior to 1994–95). An exploratory multivariate logistic regression comparing this group (regular smokers at baseline) with regular smokers from follow-up who had not reached this status at baseline was conducted. Results suggest that regular smokers at baseline may in fact be a unique group as they were found to be older, more likely to be Caucasian, reporting a lower GPA and higher rates of substance use than the those who did not reach their regular smoking status until follow-up. Our findings cannot be generalized to this group whom will require broader longitudinal inquiry.

Implications

Significant efforts in longitudinal risk research have been undertaken in recent years based on the belief that an understanding of processes involved in substance use initiation and escalation will provide the ground work for developing effective and cost-efficient prevention strategies. Although the present findings have not provided conclusive evidence regarding prevention program targets or content, they have added substantially to basic etiologic work by identifying factors that classify the majority of positive cases for experimentation and regular use, the characterization of which may provide clues into the major mechanisms at work for different groups of adolescent smokers. Although the present models did not demonstrate, as predicted, complex constellations of decision points, an outcome that in some ways would have been more consistent with our concept of "pathways," the resulting subgroups likely represent important variable combinations in need of more careful longitudinal study. Moreover, the fact that different rather than identical subgroups were identified for each stage of smoking, even when starting at the same level of use (i.e. nonsmokingp; Figs. 1 and 2) or ending with the same level of use (i.e. regular use; Figs. 2 and 3) demonstrates that our pursuit of pathways may be somewhat complicated and that our demonstration of risk patterns, possibly the first step toward defining pathways, will likely differ based on the speed with which smoking behaviors escalate. Factors that discriminate the population subgroups for the various stages of smoking should be considered in the generation of hypotheses regarding the processes that lead to substance use or its escalation.

Although this study did confirm the presence of relatively parsimonious subgroups within the population, the present models did not accurately identify many experimental and regular smokers. Ideally,

prospective assessment procedures could be developed that more sensitively identify youth that will engage in smoking while reducing the rate of false alarms. Universal prevention efforts achieve the first goal in that population-intended prevention initiatives necessarily include all future smokers (i.e. 100% sensitivity). The content of these programs, though, require substantial resources, the vast majority of which are aimed at youth that will not engage in the target behavior (i.e. 0% specificity). The emphasis on sensitivity estimates in the present study (i.e. the true positive rate) reflects the relatively acceptable tradeoff between sensitivity and specificity among prognostic tests in which false positives (i.e. youth who test positive, but do not exhibit the behavior prospectively) are more acceptable than false negatives (i.e. youth who will develop the behavior, but are missed by testing; for review of test evaluation see Kraemer, 1992).

ACKNOWLEDGMENTS

This research was supported in part by a Faculty Scholars Award, Robert Wood Johnson Foundation, Tobacco Etiology Research Network (Dr. Dierker), and uses data from the Add Health project, a program project designed by J. Richard Udry (PI) and Peter Bearman, and funded by Grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding from 17 other agencies. Persons interested in obtaining data files from The National Longitudinal Study of Adolescent Health should contact Add Health, Carolina Population Center, 123 West Franklin Street, Chapel Hill, NC 27516-2524 (http://www.cpc.unc.edu/addhealth).

REFERENCES

Adalbjarnardottir, S., & Rafnsson, F. D. (2002). Adolescent antisocial behavior and substance use: Longitudinal analyses. *Addictive Behaviors*, *27*(2), 227–240.

Alexander, C. S., Allen, P., Crawford, M. A., & McCormick, L. K. (1999). Taking a first puff: Cigarette smoking experiences among ethnically diverse adolescents. *Ethnicity and Health*, *4*(4), 245–257.

Allen, O., Page, R. M., Moore, L., & Hewitt, C. (1994). Gender differences in selected psychosocial characteristics of adolescent smokers and nonsmokers. *Health Values*, *18*(2), 34–39.

Ary, D. V., & Biglan, A. (1988). Longitudinal changes in adolescent cigarette smoking behavior: onset and cessation. *Journal of Behavioral Medicine*, *11*(4), 361–382.

Bates, M. E. (2000). Integrating person-centered and variable-centered approaches in the study of developmental courses and transitions in alcohol use: Introduction to the special section. *Alcoholism: Clinical and Experimental Research*, *24*(6), 878–881.

Botvin, G. J., Baker, E., Goldberg, C. J., Dusenbury, L., Botvin, E. (1992). Correlates and predictors of smoking among Black adolescents. *Addictive Behaviors*, *17*(2), 97–103.

Boyle, M. H., Offord, D. R., Racine, Y. A., Fleming, J. E., & Links, P. S. (1993). Predicting substance abuse in early adolescence based on parent and teacher assessments of childhood psychiatric disorder: Results from the Ontario Child Health Study follow-up. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *34*(4), 535–544.

Breiman, L., Friedman, J. H., Ohlsen, R. A., & Stone, C. J. (1984). *Classification and regression trees.* Belmont: Wadsworth International Group.

Brown, R. A., Lewinsohn, P. M., Seeley, J. R., & Wagner, E. F. (1996). Cigarette smoking, major depression, and other psychiatric disorders among adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(12), 1602–1610.

Bryant, A. L., Schulenberg, J., Bachman, J. G., O'Malley, P. M., & Johnston, L. D. (2001). Understanding the links among school misbehavior, academic achievement, and cigarette use: A national panel study of adolescents. *Prevention Science*, *1*(2), 71–87.

CDC (2002). *Annual smoking—Attributable mortality, years of potential life lost, and economic costs—United States, 1995–1999* (Morbidity and Mortality Weekly Report). Washington, DC: U.S. Department of Health and Human Services.

Chassin, L., Presson, C. C., Pitts, S. C., & Sherman, S. J. (2000). The natural history of cigarette smoking from adolescence to adulthood in a midwestern community sample: Multiple trajectories and their psychosocial correlates. *Health Psychology*, *19*(3), 223–231.

Chassin, L., Presson, C. C., Sherman, S. J., & Edwards, D. A. (1991). Four pathways to young-adult smoking status: Adolescent social–psychological antecedents in a midwestern community sample. *Health Psychology*, *10*(6), 409–418.

Chassin, L., Presson, C. C., Sherman, S. J., & Edwards, D. A. (1992). Parent educational attainment and adolescent cigarette smoking. *Journal of Substance Abuse*, *4*(3), 219–234.

Choi, W. S., Pierce, J. P., Gilpin, E. A., Farkas, A. J., & Berry, C. C. (1997). Which adolescent experimenters progress to established smoking in the United States. *American Journal of Preventive Medicine*, *13*(5), 385–391.

Cloninger, C., Bohman, M., & Sigvardsson, S. (1981). Inheritance of alcohol abuse: cross-fostering analysis of adopted men. *Archives of General Psychiatry*, *38,* 861–868.

Cloninger, C., Sigvardsson, S., Gilligan, S., von Knorring, A., Reich, T., & Bohman, M. (1988). Genetic heterogeneity and the classification of alcoholism. *Advances in Alcohol and Substance Abuse*, *7,* 3–16.

Conrad, K. M., Flay, B. R., & Hill, D. (1992). Why children start smoking cigarettes: Predictors of onset. *British Journal of Addiction*, *87*(12), 1711–1724.

Curran, G. M., Stoltenberg, S. F., Hill, E. M., Mudd, S. A., Blow, F. C., & Zucker, R. A. (1999). Gender differences in the relationship among SES, family history of alcohol disorders and alcohol dependence. *Journal of Studies on Alcohol*, *60*(6), 825–832.

Dielman, T. E., Campanelli, P. C., Shope, J. T., & Butchart, A. T. (1988). Susceptibility to peer pressure, self-esteem, and health locus of control as correlates of adolescent substance abuse. *Health Education Quarterly*, *14*(2), 207–221.

Duncan, S. C., Duncan, T. E., & Hops, H. (1998). Progressions of alcohol, cigarette, and marijuana use in adolescence. *Journal of Behavioral Medicine*, *21*(4), 375–388.

Elders, M., & Perry, C. (1994). Preventing tobacco use among young people: a report of the Surgeon General. *MMWR*, *43*, 2–10.

Epstein, J. A., Botvin, G. J., Baker, E., & Diaz, T. (1999). Impact of social influences and problem behavior on alcohol use among inner-city Hispanic and black adolescents. *Journal of Studies on Alcohol*, *60*(5), 595–604.

Feigelman, W., & Lee, J. (1995). Probing the paradoxical pattern of cigarette smoking among African-Americans: Low teenage consumption and high adult use. *Journal of Drug Education*, *25*(4), 307–320.

Flay, B. R., Hu, F. B., Siddiqui, O., Day, L. E., *et al.* (1994). Differential influence of parental smoking and friends' smoking on adolescent initiation and escalation of smoking. *Journal of Health and Social Behavior*, *35*(3), 248–265.

Flint, A. J., Yamada, E. G., & Novotny, T. E. (1998). Black-white differences in cigarette smoking uptake: Progression from adolescent experimentation to regular use. *Preventive Medicine: An International Devoted to Practice and Theory*, *27*(3), 358–364.

Forster, J. L., Wolfson, M., Muray, D. M., Wagenaar, A. C., & Claxton, A. J. (1997). Perceived and measured availability of tobacco to youths in 14 Minnesota communities: The TPOP study. *American Journal of Preventive Medicine*, *13*(3), 167–174.

Glantz, M. D. (1992). A developmental psychopathology model of drug abuse vulnerability. In Glantz, M., Pickens, R. W. (Eds.), *Vulnerability to drug abuse* (pp. 389–418). Washington, DC: American Psychological Association.

Glantz, M. D., & Pickens, R. W. (Eds.) (1992). *Vulnerability to drug abuse*. Washington, DC: American Psychological Association Press.

Griesler, P. C., & Kandel, D. B. (1998). Ethnic differences in correlates of adolescent cigarette smoking. *Journal of Adolescent Health*, *23*(3), 167–180.

Gritz, E. R., Prokhorov, A. V., Hudmon, K. S., Chamberlain, R. M., Taylor, W. C., DiClemente, C. C., Johnston, D. A., Hu, S., Jones, L. A., Jones, M. M., Rosenblum, C. K., Ayars, C. L., & Amos, C. I. (1998). Cigarette smoking in a multiethnic population of youth: Methods and baseline findings. *Preventive Medicine: An International Devoted to Practice and Theory*, *27*(3), 365–384.

Group, Conduct Problems Prevention Research (2000). Merging universal and indicated prevention programs: The Fast Track model. *Addictive Behaviors*, *25,* 913–927.

Harrell, J. S., Bangdiwala, S. I., Deng, S., Webb, J. P., & Bradley, C. (1998). Smoking initiation in youth: The roles of gender, race, socioeconomics, and developmental status. *Journal of Adolescent Health*, *23*(5), 271–279.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992, July). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin*, *112*(1), 64–105.

Headen, S. W., Bauman, K. E., Deane, G. D., & Koch, G. G. (1991). Are the correlates of cigarette smoking initiation different for Black and White adolescents? *American Journal of Public Health*, *81*(7), 854–858.

IOM (1996). *A plan for prevention research for the NIMH* (NIH 96-4093). Washington, DC: National Institute of Mental Health.

Jackson, C. (1997). Initial and experimental stages of tobacco and alcohol use during late childhood: Relation to peer, parent, and personal risk factors. *Addictive Behaviors*, *22*(5), 685–698.

Jackson, K. M., Sher, K. J., Cooper, M. L., & Wood, P. K. (2002). Adolescent alcohol and tobacco use: Onset, persistence and trajectories of use across two samples. *Addiction*, *97*(5), 517–531.

Jessor, R., & Jessor, L. (1977). *Problem behavior and psychosocial development—A longitudinal study of youth*. New York: Academic Press.

Juon, H.-S., Ensminger, M. E., & Sydnor, K. D. (2002). A longitudinal study of developmental trajectories to young adult cigarette smoking. *Drug and Alcohol Dependence*, *66*(3), 303–314.

Kandel, D. B. (1982). Epidemiological and psychosocial perspectives on adolescent drug use. *Journal of the American Academy of Child Psychiatry*, *21*(4), 328–347.

Kandel, D. B., Wu, P., & Davies, M. (1994). Maternal smoking during pregnancy and smoking by adolescent daughters. *American Journal of Public Health*, *84*(9), 1407–1413.

Kraemer, H. C. (1992). *Evaluating medical tests*. Newbury Park, CA: Sage.

Kumpfer, K. L., Molgaard, V., & Spoth, R. (1996). The Strengthening Families Program for the prevention of delinquency and drug use. In R. J. McMahon (Ed.), *Preventing childhood disorders, substance abuse, and delinquency* (pp. 241–267). Thousand Oaks, CA: Sage.

Landrine, H., Richardson, J. L., Klonoff, E. A., & Flay, B. R. (1994). Cultural diversity in the predictors of adolescent cigarette smoking: The relative influence of peers. *Journal of Behavioral Medicine*, *17*(3), 331–346.

Mackesy-Amiti, M. E., Fendrich, M., & Goldstein, P. J. (1997). Sequence of drug use among serious drug users: Typical vs. atypical progression. *Drug and Alcohol Dependence*, *45*(3), 185–196.

Maddahian, E., Newcomb, M. D., & Bentler, P. M. (1987). Adolescents' substance use: Impact of ethnicity, income, and availability. *Advances in Alcohol and Substance Abuse*, *5*(3), 63–78.

Maddahian, E., Newcomb, M. D., & Bentler, P. M. (1988). Risk factors for substance use: Ethnic differences among adolescents. *Journal of Substance Abuse*, *1*(1), 11–23.

Mayhew, K., Flay, B. R., & Mott, J. A. (2000). Stages in the development of adolescent smoking. *Drug and Alcohol Dependence, 59*(Suppl. 1), S61–S81.

McMahon, R. J., & Slough, N. M. (1996). Family-based intervention in the Fast Track Program. In R. J. McMahon (Ed.), *Preventing childhood disorders, substance abuse, and delinquency* (pp. 90–110). Thousand Oaks, CA: Sage.

Miller-Johnson, S., Lochman, J. E., Coie, J. D., Terry, R., & Hyman, C. (1998). Comorbidity of conduct and depressive problems at sixth grade: Substance use outcomes across adolescence. *Journal of Abnormal Child Psychology*, *26*(3), 221–232.

Mittelmark, M. B., Murray, D. M., Luepker, R. V., Pechacek, T. F., Pirie, P. L., & Pallonen, U. E. (1987). Predicting experimentation with cigarettes: The childhood antecedents of smoking study (CASS). *American Journal of Public Health*, *77*(2), 206–208.

Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

Oetting, E. R., & Beauvais, F. (1986). Peer cluster theory: Drugs and the adolescent. *Journal of Counseling and Development*, *65*(1), 17–22.

Oetting, E. R., & Beauvais, F. (1987). Peer cluster theory, socialization characteristics, and adolescent drug use: A path analysis. *Journal of Counseling Psychology*, *34*(2), 205–213.

Patton, G. C., Carlin, J. B., Coffey, C., Wolfe, R., Hibbert, M., & Bowes, G. (1998). Depression, anxiety, and smoking initiation: A prospective study over 3 years. *American Journal of Public Health*, *88*(10), 1518–1522.

Pederson, L. L., Koval, J. J., McGrady, G. A., & Tyas, S. L. (1998). The degree and type of relationship between psychosocial variables and smoking status for students in Grade 8: Is there a dose-response relationship? *Preventive Medicine: An International Devoted to Practice and Theory*, *27*(3), 337–347.

Pierce, J. P., Choi, W. S., Gilpin, E. A., Farkas, A. J., & Merritt, R. K. (1996). Validation of susceptibility as a predictor of which adolescents take up smoking in the United States. *Health Psychology*, *15*(5), 355–361.

Pizzolongo, P. J. (1996). The Comprehensive Child Development Program and other early intervention program models. In

R. J. McMahon (Ed.), *Preventing childhood disorders, substance abuse, and delinquency* (pp. 48–64). Thousand Oaks, CA: Sage.

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*(3), 390–395.

Program, The Fast Track (1992). A developmental and clinical model for the prevention of conduct disorder. *Development and Psychopathology*, *4*(4), 509–527.

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., Tabor, J., Beuhring, T., Sieving, R. E., Shew, M., Ireland, M., Bearinger, L. H., & Udry, J. R. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *JAMA*, *278*(10), 823–832.

Robinson, L. A., Klesges, R. C., Zbikowski, S. M., & Glaser, R. (1997). Predictors of risk for different stages of adolescent smoking in a biracial sample. *Journal of Consulting and Clinical Psychology*, *65*(4), 653–662.

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1991, May). Comorbidity of unipolar depression: II. Comorbidity with other mental disorders in adolescents and adults. *Journal of Abnormal Psychology*, *100*(2), 214–222.

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1996). Psychiatric comorbidity with problematic alcohol use in high school students. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(1), 101–109.

Sieber, M. F., & Angst, J. (1990). Alcohol, tobacco and cannabis: 12-year longitudinal associations with antecedent social context and personality. *Drug and Alcohol Dependence*, *25*(3), 281–292.

Simon, T. R., Sussman, S., Dent, C. W., Burton, D., *et al.* (1995). Prospective correlates of exclusive or combined adolescent use of cigarettes and smokeless tobacco: A replication-extension. *Addictive Behaviors*, *20*(4), 517–524.

Steinberg, D., & Colla, P. (1995). *CART: Tree-structured nonparametric data analysis.* San Diego: Salford Systems.

Sussman, S., Dent, C. W., & Galaif, E. R. (1997). The correlates of substance abuse and dependence among adolescents at high risk for drug abuse. *Journal of Substance Abuse*, *9,* 241–255.

Tyas, S. L., & Pederson, L. L. (1998). Psychosocial factors related to adolescent smoking: A critical review of the literature. *Tob Control*, *7*(4), 409–420.

USDHHS (1994). *Preventing tobacco use among young people: A report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

USDHHS (1998). *Tobacco use among U.S. racial/ethnic minority groups: A Report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Van Roosmalen, E. H., & McDaniel, S. A. (1992). Adolescent smoking intentions: Gender differences in peer context. *Adolescence*, *27*(105), 87–105.

Westfall, P. H., Tobias, R. D., Rom, D., Wolfinger, R. D., & Hochberg, Y. (1999). *Multiple comparisons and multiple tests using the SAS(R) system*. Cary, NC: SAS Institute.

Zhang, H. P., & Bracken, M. B. (1996). A tree-based two-stage risk factor analysis of spontaneous abortion. *American Journal of Epidemiology*, *144,* 989–996.

Zucker, R. (1994). Pathways to alcohol problems and alcoholism: A developmental account of the evidence for multiple alcoholisms and for contextual contributions to risk. In G. Boyd (Ed.), *The development of alcohol problems: Exploring the biopsychosocial matrix of risk* (pp. 255–289). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.