Global Social Welfare
https://doi.org/10.1007/s40609-021-00220-9

ORIGINAL ARTICLE

Temporal Patterns of Cigarette Smoking and Its Associated Covariates: a Multilevel Longitudinal Data Analysis

David Adzrago[1] · Lucy Kavi[2] · Rosemary I. Ezeugoh[2] · Benedicta Osafo-Darko[3]

Accepted: 10 December 2021
This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021

Abstract

Background Understanding differences in cigarette smoking patterns such as the frequency between-person and within-person is essential for tailored tobacco health education interventions. Previous studies, however, mostly limited analysis to computation of cigarette smoking frequency and its correlates. This article used multilevel models to examine between-person and within-person variations in cigarette smoking patterns over a 13-year period.

Methods We merged the National Longitudinal Study of Adolescent Health public-use data waves 1–4 into one longitudinal dataset for use in this study. Our analysis was based on the past-month's average number of cigarette smoked per day. We used linear mixed model approach to fit multilevel models.

Results The average number of cigarette smoked per day (CPD) among the sample at baseline/wave 1 was 6.92 (SD = 8.18). Time of observation in years ($\beta = 0.455$ ($p < .001$), age ($\beta = 0.355$, $p < .001$), past-year alcohol use frequency ($\beta = -0.329$, $p < .001$), and illicit drug use ($\beta = 1.128$, $p < .001$) were associated with average number of CPD. There were significant variations in the average number of CPD between-person ($\beta = 29.602$, $p < .001$) and within-person (variance = 34.393, $p < .001$).

Conclusions This study demonstrates that rate of change in average number of CPD over years among the study sample could be different between-adolescent and within-adolescent depending on other substance use and demographic factors. Hence, tailored tobacco use educational programs or interventions and policies targeting these adolescents could be designed according to between-adolescent and within-adolescent differences in the average number of CPD trajectories.

Keywords Cigarette smoking · Marijuana use · Alcohol use · Illicit drug use

Introduction

Tobacco use continues to be a significant public health problem in the United States (US). It contributes to over 480,000 deaths annually and remains the most significant preventable cause of disease and disability in the US (King et al., 2018; Sung et al., 2018; US Department of Health & Human Services, 2014). Cigarette smoking is a major cardiovascular disease risk factor and contributes approximately 90% of lung cancer prevalence in the US (Alberg et al., 2013). Additionally, cigarette smoking increases the risks of financial burdens and loss of productivity for cigarette smokers (Alberg et al., 2003; Hu et al., 2020; Thompson et al., 2015). Despite the health and economic risks associated with cigarette smoking, daily cigarette smoking at an early age persists at high rates in the US (Bonnie et al., 2015; Cantrell et al., 2018; US Department of Health & Human Services, 2012). In 2012, it was estimated that most adult daily cigarette smokers had started smoking cigarettes before 18 years (87%) and 21 years (95%) (US Department of Health & Human Services, 2012). Considering the early age at which daily cigarette smoking behavior is initiated, the health and economic consequences of cigarette smoking may get worse as adolescents develop into adulthood. Thus, cigarette smoking behavior needs to be evaluated longitudinally to determine its changing patterns over time and age.

Highlights
- This is the first study to examine between-person and within-person variations in the number of cigarette smoked per day over a 13-year period.
- The rate of change in the average number of cigarette smoked per day was associated with every year increase in the observation period.
- Age was associated with increased average number of number of cigarette smoked per day.
- Illicit drug use was associated with increased average number of cigarette smoked per day.
- Average number of cigarette smoked per day differed among the sample due to between-person and within-person differences.

✉ David Adzrago
David.Adzrago@uth.tmc.edu

Extended author information available on the last page of the article

Cigarette smoking among adolescents remains a public health concern because a substantial proportion of them still smoke cigarettes (Johnston et al., 2019). Results from previous studies that have evaluated the decline in daily cigarette smoking in this population have been contradictory. Johnston et al. (2019), for instance, determined a marked decline in cigarette smoking prevalence from approximately 40% among 12th graders in 1995 to 10% in 2018 (Johnston et al., 2019). However, another study found that the overall decline in daily cigarette use was not significant among adolescents (Jamal et al., 2017). Again, previous longitudinal studies revealed that daily cigarette smoking may progress from low daily cigarette smoking in adolescence to high daily cigarette smoking in adulthood (Chassin et al., 1990; Choi et al., 1997; Hu et al., 2020; Sargent et al., 2017). Most of these longitudinal studies (Chassin et al., 1990; Choi et al., 1997; Sargent et al., 2017), however, were only conducted over short periods (2–8 years) occluding the influence of a longer timeframe on daily cigarette use. These studies revealed that gender, race/ethnicity, aggressive behavior (e.g., damaging properties), and the use of other substances such as alcohol, marijuana, and other illicit drugs could potentially explain the patterns of cigarette smoking from adolescence to adulthood (Chassin et al., 1990; Choi et al., 1997; Hu et al., 2020; Sargent et al., 2017).

Furthermore, previous studies (Chassin et al., 1990; Choi et al., 1997; Goldade et al., 2012; Sargent et al., 2017) including longitudinal studies largely examined or evaluated cigarette smoking frequency and it correlates without assessing the differences in cigarette smoking frequency between-person and within-person. These within-person and between-person variations in cigarette smoking patterns are important as well because they can help to effectively tailor interventions at individual and national levels (Maruyama et al., 2009; Piasecki et al., 2016; Singer, 1998). Multilevel models are instrumental in examining between-person and within-person variations in cigarette smoking patterns, where repeated measures of cigarette smoking frequency are nested within persons (Belsky et al., 2013; Piasecki et al., 2016; Singer, 1998). Thus, these models help to examine changes in cigarette smoking frequency among persons as a function of time. However, there is a paucity of studies that examined between-person and within-person variations in cigarette smoking frequency using longitudinal data.

To address this gap/limitation, our current study used secondary data extracted from the National Longitudinal Study of Adolescent Health (Add Health) public-use data waves 1–4 to examine changes in the average number of cigarettes smoked each day (CPD) over a 13-year period. Our objective was to determine the degrees of within-person and between-person variations in the average number of CPD over a 13-year period. We also assessed the association between the average number of CPD and other substance use frequencies and aggressive behavior, respectively.

Methods

Study Sample

We extracted secondary data from the National Longitudinal Study of Adolescent Health (Add Health) public-use data waves 1–4. The Add Health study surveyed a nationally representative sample of US adolescents who were followed as they transitioned into emerging adulthood (see details in Chen & Chantala, 2014; Harris et al., 2009). The survey was administered as an in-school questionnaire in the 1994–1995 school year and followed by a series of in-home interviews in 1995, 1996, 2001–2002, 2008, and 2016–2018. The in-home interviews conducted represent five waves, respectively. The adolescents recruited were in grades 7 to 12 (aged 12–19 years) at wave I, grades 8 to 12 at wave II, aged 18 to 26 years at wave III, aged 24 to 32 years at wave IV, and aged 32 to 42 at wave V. The Add Health study sample was selected from a sample of 80 high schools and 52 middle schools. The schools were selected with unequal probability. To ensure that the sample was representative of US schools concerning the region of the country, urbanicity, school size, school type, and ethnicity, systematic sampling methods and implicit stratification were incorporated into the Add Health study design. The Institutional Review Board (IRB) at the University of North Carolina at Chapel Hill, North Carolina approved the Add Health study protocol. The data from the in-home interviews at waves I–IV were used for this present study. We used the public-use dataset (wave I = 6504, wave II = 4834, wave III = 4882, and wave IV = 5114) from the in-home interviews. These four-wave datasets were merged to evaluate our research objectives. The wave V data was not added because some of the variables we included in our study were not available or the same in the wave V dataset. For instance, while marijuana use frequency was measured in waves 1–4 as the number of times in the past 30 days, it was measured in wave V as the number of days in the past 30 days. Also, school attendance (If SCHOOL YEAR: Are you presently in school? If SUMMER: Were you in school during this past school year?) question was not asked at wave V.

Measures

Outcome

Past-Month Cigarette Use Per Day.

This variable was measured at all the four waves by asking the participants, "During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?" We analyzed this variable as a continuous variable. Hence, the unit of analysis is adolescents who ever smoked cigarette during the past 30 days.

Predictors

Time in Years We constructed this variable to indicate the time in years that the data were collected from the study participants at each wave or observation (Time = 0 [1995], Time = 1 [1996], Time = 6 [2001], and Time = 13 [2008]). The following predictors were measured at all waves.

Marijuana Use Frequency Marijuana use was measured by asking the participants to indicate, "During the past 30 days, how many times did you use marijuana?".

Alcohol Use Frequency This variable was also measured as, "During the past 12 months, on how many days did you drink alcohol?".

Illicit Drug Use The participants were asked "How old were you when you first tried any other type of illicit drugs, such as LSD, PCP, ecstasy, mushrooms, speed, ice, heroin, or pills, without a doctor's prescription? If you never tried any other type of illicit drug, enter "0." This variable was dichotomized (tried any other illicit drug = 1 and never tried any illicit drug = 0).

Aggressive Behavior This variable was determined by asking, "In the past 12 months, how often did you deliberately damage property that didn't belong to you?".

Covariates

School attendance (If SCHOOL YEAR: Are you presently in school? If SUMMER: Were you in school during this past school year? Yes = 1 and no = 0), age, gender (male or female), and race (Hispanic or Latino = 1, non-Hispanic Black or African American = 2, non-Hispanic Asian or Pacific Islander = 3, non-Hispanic American Indian or Native American = 4, non-Hispanic Other = 5, and non-Hispanic White) were controlled for.

Statistical Analyses

Our analysis was based on the sample that had complete data on the average number of CPD. We performed an unweighted analysis because we did not intend to generalize our findings to the entire US adolescent and emerging adult populations. Our findings would be generalizable only to the sample we analyzed. Multicollinearity was examined using variance inflation factor (VIF) and tolerance (Tol) to determine the degree of collinearity between the independent variables or predictors in the data (Alin, 2010; Kim, 2019; Thompson et al., 2017; Yoo et al., 2014). Based on existing literature, we considered tolerance values below 0.1 (cutoff) and VIF values greater than 10 (cutoff) as suggestive of multicollinearity (Alin, 2010; Kim, 2019; Thompson et al., 2017; Yoo et al., 2014). There appeared to be no issues or threats of multicollinearity based on these criteria since all tolerance and VIF values were below 0.1 and 10 respectively. We calculated descriptive statistics of participant characteristics and presented the summaries in Table 1. Multilevel models were used to evaluate the degree of within-person and between-person variations in the average number of CPD over the 13-year period. We specifically examined changes in the average number of CPD among persons as a function of time (i.e., 13 years). These models were applied because the repeated measures of the average number of CPD were nested within persons.

Data analyses were performed using SAS 9.4 (SAS, 2013), and an alpha level of 0.05 was used to determine significant estimates. We used PROC MIXED, a linear mixed model (LMM) procedure in SAS, to fit the models (Piasecki et al., 2016; Singer, 1998). PROC MIXED was used to model the random and fixed effects as well as accommodate missing data, which are common in longitudinal studies (Piasecki et al., 2016; Singer, 1998). The LMMs were applied to examine the continuous outcome, the average number of CPD. The multilevel models in our analysis included a two-level model: level-1 model and level-2 model. Our level-1 model determined the linear model, and the level-2 model examined the variations in parameters from the linear model as random effects (Maruyama et al., 2009; Singer, 1998). We first fitted an unconditional means model or unconditional linear model (i.e., the model with no level-2 or person-level predictors) by evaluating the two levels (i.e., level 1 = time points and level 2 = person) with the average number of CPD. In all our model estimations of the random effects, we allowed the variance and covariance to be correlated within-persons and within-time years/waves. We then calculated intraclass correlation coefficient (ICC) to determine the total variance in the average number of CPD that was attributable to between-person and intra-person differences (Piasecki et al., 2016; Singer, 1998). The ICC, which determines the total variance due to between-person differences, was calculated as follows (Piasecki et al., 2016; Singer, 1998): $ICC = (\sigma^2_U)/(\sigma^2_U + \sigma^2_e) = (28.4295)/(28.4295 + 39.0794) = 0.4211$. This means that 42.11% of the total variance in the average number of CPD is attributable to between-person differences. However, majority of

Table 1 Participant characteristics at baseline/wave 1 ($N = 6504$)

Variable	Summary statistics		
	n	Mean (%)	SD
Age	6504	15.53	1.78
Gender			
Males	3147	48.39%	–
Females	3356	51.61%	–
Race			
Non-Hispanic Whites	3736	57.50%	–
Non-Hispanic Blacks or African Americans	1584	24.38%	–
Non-Hispanic Asian or Pacific Islander	247	3.80%	–
Non-Hispanic American Indian/Native American	131	2.02%	–
Hispanics or Latinos	743	11.44%	–
Non-Hispanic others	56	0.86%	–
School attendance			
Yes	6376	98.08%	–
No	125	1.92%	–
Past-month CPD	1653	6.92	8.18
Past-year alcohol use frequency	3545	3.92	2.12
Past-month marijuana use frequency	1655	7.22	35.20
Illicit drug use			
Ever used	499	7.79%	–
Never used	5903	92.21%	–
Past-year aggressive behavior frequency			
Never	5303	82.31%	–
1 or 2 times	904	14.03%	–
3 or 4 times	135	2.10%	–
5 or more times	101	1.57%	–

CPD cigarette smoked per day, *SD* standard deviation, % percentage

the total variance (57.89%) was due to intra-person differences. Our result showed significant variation in the average number of CPD between-person ($\beta = 28.4295$, $p < 0.001$) and within-person (variance = 39.0794, SE = 1.1553, Wald $z = 33.83$, $p < 0.001$). The slope for the linear time trend was significant ($\beta = 0.8259$, $p < 0.001$), which further confirms the fitting of the LMM.

The results further suggest that level 2 (person-level) covariates could potentially explain the variance in the average number of CPD between-person and within-person. We fitted a conditional linear model with gender, a person-level covariate, after fitting the unconditional model. This was to determine whether the variations in the intercepts and slopes could be due to a covariate. We compared the Bayesian information criterion (BIC) and Akaike information criterion (AIC) of the unconditional and conditional models to select the model that fits better. The model with the lower BIC and AIC suggests a better fit (Liu et al., 2012).

The BIC (46,736.9) and AIC (46,709.7) for the unconditional and BIC (46,708.9) and AIC (46,681.8) for the conditional model show that the conditional model fits better. The conditional model demonstrated the following results. The fixed effects showed that the average number of CPD is 5.7235 ($p < 0.001$) across persons at baseline (i.e., time = 0), adjusting for gender and interaction between gender and time. The growth rate was statistically significant ($\beta = 0.5814$, $p < 0.001$), adjusting for gender and its interaction with time. Gender was associated with the average number of CPD ($\beta = 0.9273$, $p = .004$). There was no significant interaction between gender and time ($\beta = 0.1301$, $p = 0.062$). The random effects showed that between-person variance has significantly reduced from 28.4295 to 28.1396 while within-person variance has increased marginally from 39.0794 to 39.1093. Computing this change in variations between-person (28.4295–28.1396)/28.4295 = 0.0102, suggests that gender accounts for about 1.02% reduction in the variance in the average number of CPD between-person. Computing the increase in variance within-person (39.1093–39.0794)/39.1093 = 0.0008, suggests that gender accounts for about 0.08% increase in the variance in the average number of CPD within-person. The variance component for the growth rates went from 0.8259 to 0.8145. Computing (0.8259–0.8145)/0.8259 = 0.0138, we find a 1.38% reduction. Thus, gender accounts for 1.38% of the explainable variations in

the average number of CPD growth rates. Hence, our final model (i.e., conditional model) included additional person-level covariates that could potentially explain the variations in the average number of CPD (see Table 2). We did not, however, include the interaction term for gender and time because it was not statistically significant.

Results

Table 1 shows that the average number of past-month CPD at baseline/wave 1 was 6.92 with high variabilities between participants (SD = 8.18). The participants had an average age of 15.53 years old (SD = 1.78). The majority of the participants were females (51.61%), non-Hispanic Whites (57.50%), and were currently in school as of the interview year or in school during the past school year (98.08%). The participants on average used alcohol in the past year for 3.92 days (SD = 2.12) and used marijuana on average 7.22 times in the past month (SD = 35.20). An important proportion of the participants ever used illicit drugs (7.79%) and engaged in aggressive/delinquent behavior one or two times in the past year (14.03%).

For every variable, the rest of variables in the model were adjusted for. The fixed effects in Table 2 show that for every increase in time of observation in years (i.e., time in years), the estimated average number of CPD across persons statistically increased by a factor of 0.455 ($p < 0.001$). Age is significantly associated with the average number of CPD ($\beta = 0.355$, $p < 0.001$). Thus, for every 1-year increase in age, the average number of CPD significantly increased by a factor of 0.355. Compared to females, the estimated

Table 2 Factors associated with between-person and within-person variations in the average number of CPD over 13 years of observation period

Fixed effects	*β*	SE	DF	t value	*P*
Intercept	3.858	1.911	2362	2.02	**0.044**
Time in years	0.455	0.054	1006	8.48	**< 0.001**
Age	0.355	0.092	380	3.87	**< 0.001**
Gender					
Female	Ref	–	–	–	–
Male	1.788	0.304	380	5.89	**< 0.001**
Race					
Non-Hispanic Whites	Ref	–	–	–	–
Non-Hispanic Blacks or African Americans	−4.421	0.430	380	−10.27	**< 0.001**
Non-Hispanic Asian or Pacific Islanders	−4.324	0.889	380	−4.87	**< 0.001**
Non-Hispanic American Indian/Native Americans	2.215	0.937	380	2.36	**0.019**
Hispanics or Latinos	−3.235	0.529	380	−6.11	**< 0.001**
Non-Hispanic other	−3.737	1.808	380	−2.07	**0.039**
School attendance					
No	–	–	–	–	–
Yes	−0.928	0.971	380	−0.96	0.340
Past-year alcohol use	−0.329	0.086	380	−3.84	**< 0.001**
Past-month marijuana use frequency	−0.006	0.004	380	−1.73	0.085
Illicit drug use					
Never used	Ref	–	–	–	–
Ever used	1.128	0.282	380	4.01	**< 0.001**
Past-year aggressive behavior frequency					
Never	Ref	–	–	–	–
1 or 2 times	−0.386	0.367	380	−1.05	0.293
3 or 4 times	1.534	0.809	380	1.89	0.059
5 or more times	1.414	0.944	380	1.50	0.135
Random effects (covariance parameter estimates)		**β**	**SE**	**Z value**	***P***
Intercept (between-person variance)		29.602	2.882	10.27	**< 0.001**
Covariance between time in years and intercept		−2.150	0.491	−4.38	**< 0.001**
Time in years		0.776	0.124	6.27	**< 0.001**
Residual (Within-person variance)		34.393	1.654	20.79	**< 0.001**

DF degree of freedom, *Ref* reference group, *SE* standard error, *β* beta

average number of CPD statistically increased by a factor of 1.788 ($p < 0.001$) for males. Non-Hispanic Blacks or African Americans ($\beta = -4.421$, $p < 0.001$), non-Hispanic Asian/Pacific Islanders ($\beta = -4.324$, $p < 0.001$), Hispanics or Latinos ($\beta = -3.235$, $p < 0.001$), and non-Hispanic others ($\beta = -3.737$, $p = 0.039$) had significantly lower average number of CPD than non-Hispanic Whites. For every additional day in past-year alcohol use, the average number of CPD significantly decreased by a factor of 0.329 ($p < 0.001$). Compared to the participants who had never used illicit drugs, the estimated average number of CPD increased significantly by a factor of 1.128 ($p < 0.001$) those who ever used illicit drugs.

The random effects showed significant variations in the average number of CPD between-person ($b = 29.602$, $p < 0.001$) and within-person (variance = 34.393, $p < 0.001$). Computing $(29.602)/(29.602 + 34.393) = 0.4626$ suggests that about 46.26% of the total variance in the average number of CPD is attributable to between-person differences conditional on the covariates. However, 53.74% of the total variance was due to intra-person differences conditional on the covariates. The results further indicate that there is a significant correlation between the intercept and the slope (covariance component; $\beta = -2.150$, $p < 0.001$). Thus, the effect of time in years on the average number of CPD differs depending on the average number of CPD within-person. Additionally, the results showed that the rate of change in the average number of CPD was 0.776 ($p < 0.001$) for every year increase in the observation period.

Discussion

This study conducted a secondary analysis on national longitudinal survey data to examine changes in the average number of cigarettes smoked each day (CPD) over a 13-year period. Unlike prior studies, this study further examined the degrees of within-person and between-person variations in the average number of CPD over 13 years of the observation period. Furthermore, we assessed the association between the average number of CPD and marijuana use frequency, alcohol use frequency, illicit drug use, and aggressive behavior. We applied multilevel models in PROC MIXED to evaluate our objectives. Thus, we adopted linear mixed models (LMMs) to model the average number of CPD patterns over 13-year observation period.

Our results show that the average number of CPD over the 13 years of observation, in general, was high among the participants. The results further suggested that as time in years increased, the average number of CPD increased. A possible reason to these findings is that the adolescents could become dependent on CPD as the number of CPD increases, which would have also progressed over time if potential tobacco cessation interventions are not implemented for them. These results are consistent with previous studies that reported a similar increase in CPD (Bonnie et al., 2015; Chassin et al., 2008; Hu et al., 2020; O'Loughlin et al., 2014; Saddleson et al., 2016). Comparable to previous longitudinal studies (Chassin et al., 1990; Choi et al., 1997; Hu et al., 2020; Sargent et al., 2017), this current study found age to be associated with an increase in the average number of CPD. This suggests that the average number of CPD may progress from low to high from adolescence to adulthood. This may reflect CPD dependency dynamics among the adolescents and adults because the transition from adolescence to adulthood is also accompanied by increase in CPD. Similarly, our result is consistent with previous studies (Jamal et al., 2017; O'Loughlin et al., 2014; Thompson et al., 2015; US Department of Health & Human Services, 2012) that determined that gender and race were associated with the average number of CPD. Our results reveal that although our sample comprised fewer males than females, the average number of CPD was significantly higher in males. Additionally, we determined that non-Hispanic Blacks, non-Hispanic Asian or Pacific Islanders, Hispanics or Latinos, and non-Hispanic others had a lower average number of CPD compared to non-Hispanic Whites. This was, again, consistent with results of previous studies about higher cigarette use among non-Hispanic Whites than racial/ethnic minorities (Griesler et al., 2002; Kandel et al., 2004; Sutter et al., 2018; Trinidad et al., 2004). Perhaps the non-Hispanic White adolescents tend to have more access to tobacco products, which might have resulted in increased average CPD and further progressed into their adulthood compared to other racial/ethnic minority adolescents. The racial/ethnic disparities in CPD may be due to the fact that racial/ethnic minority adolescents tend to have limited access to socioeconomic resources such as financial resources to afford more cigarettes compared to non-Hispanic White adolescents, who experience less racial discrimination, poverty, neighborhood segregation, and structural barriers (Alexander et al., 2016; Brown et al., 2014; Fagan et al., 2007; Garrett et al., 2014). This could also explain the racial/ethnic tobacco smoking paradox where non-Hispanic White adolescents have less tobacco-related diseases although they have higher cigarette smoking risks compared to racial/ethnic minority adolescents (Alexander et al., 2016).

We found that alcohol use frequency was associated with CPD, but this association was negative suggesting that adolescents may be using alcohol to reduce CPD. This also suggests that there is a need to tailor alcohol-related interventions to alcohol use among adolescents as their alcohol use may increase. The role of alcohol use in CPD is confirmed by existing literature that determined alcohol use

as a significant predictor of daily cigarette smoking (Brook et al., 2007; O'Loughlin et al., 2014; Tanner et al., 2015). Consistent with a longitudinal study by Kandel et al. (2004) that focused on CPD, our results show that aggressive/delinquent behavior was not associated with the average number of CPD. This current study additionally found that illicit drug use was positively associated with the average number of CPD. Hence, adolescents using illicit drugs were at higher risks of engaging in daily cigarette smoking than those who did not use any illicit drugs (Brook et al., 2007). Possibly, the adolescents using illicit drugs may also be engaging in other substances such as cigarette products, which may suggest that they are engaging in poly substance use.

To the best of our knowledge, no previous studies applied multilevel models to examine between-person and within-person variations in the average number of CPD over a 13-year period. The current study applied these models to bridge the gap in the literature. Our results indicate that between-person and within-person variations exist in the average number of CPD based on our covariates used. Based on these covariates, we found that approximately 46.26% of the total variations in the average number of CPD could be attributable to between-person differences. This implies that significant factors such as time in years, age, gender, race, alcohol use frequency, and illicit drug use could potentially explain the average number of CPD trajectories between adolescents. However, 53.74% of the total variations in the average number of CPD could be due to intra-person differences based on these covariates. Additionally, our findings revealed that the effect of time in years on the average number of CPD varied depending on the average number of CPD within-person conditional on the covariates. Moreover, our results showed that the rate of change in the average number of CPD was significantly positive with an estimated yearly increase in the average CPD by a factor of 0.776 within the observation period. Thus, the rate of change for a person with an average CPD of 29.602 at baseline/wave 1 will increase by a factor of 0.776 to 52.595 average CPD per observation occasion or wave on average.

This current study has some limitations that need to be noted. First, we performed an unweighted data analysis. Using national data without the sampling weights and clusters may lead to underestimating standard errors and statistical test results (Chen & Chantala, 2014). The weight and cluster adjust for estimates regarding clustering and the unequal probability of selection (Chen & Chantala, 2014). As we mentioned earlier, we did not intend to generalize our findings to the general population or make any inferences beyond the sample we analyzed. Besides, no Add Health guidelines exist on how to incorporate weight components for the type of analysis we performed at present. Furthermore, recall and respondent biases may influence our findings because the Add Health data contain self-reported information. Additionally, the data presented were not collected at equal intervals over the waves; some waves were shorter than others. Despite these limitations, our study makes a significant contribution to the literature on between-person and within-person variations in the average number of CPD over a long period of observations.

Conclusion

Our study revealed some important findings. First, differences in the average number of CPD trajectories between adolescents in this study could be explained by factors such as time in years, age, gender, race, alcohol use frequency, and illicit drug use. Based on our covariates selected, we determined some major intra-adolescent differences in the average number of CPD patterns. However, the effect of time in years on the average number of CPD varied depending on the average number of CPD within adolescents. Additionally, the rate of change in the average number of CPD over observation occasion or wave was positively significant. Thus, the average number of CPD could increase from adolescence to adulthood but the growth rates could be different between adolescents and within adolescents depending on other substance use and demographic factors. This current study demonstrates that tailored tobacco use educational programs or interventions and policies targeting these adolescents could be designed according to between-adolescent and within-adolescent differences in the average number of CPD trajectories.

Author Contribution David Adzrago served as the leading author, conducted the literature review, performed the statistical analyses, drafted the manuscript, and coordinated writing the manuscript. Lucy Kavi coordinated the manuscript writing and provided critical revisions of the manuscript. Rosemary I. Ezeugoh and Bennie Osafo-Darko provided critical revisions of the manuscript.

Availability of Data and Material Publicly available data and material at https://addhealth.cpc.unc.edu/data/

Declarations

Ethics Approval This paper was performed using de-identified public use data and therefore a review from the authors' Institutional Review Board was not required.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.

References

Alexander, L. A., Trinidad, D. R., Sakuma, K. L. K., Pokhrel, P., Herzog, T. A., Clanton, M. S., ... & Fagan, P. (2016). Why we must continue to investigate menthol's role in the African American smoking paradox. *Nicotine & Tobacco Research*, *18*(suppl_1), S91-S101.

Alberg, A. J., Brock, M. V., Ford, J. G., Samet, J. M., & Spivack, S. D. (2013). Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest, 143*(5 Suppl), e1S-e29S. https://doi.org/10.1378/chest.12-2345

Alberg, A. J., Diette, G. B., & Ford, J. G. (2003). Invited commentary: attendance and absence as markers of health status—the example of active and passive cigarette smoking.

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(3), 370–374.

Belsky, D. W., Moffitt, T. E., Baker, T. B., Biddle, A. K., Evans, J. P., Harrington, H., ... & Poulton, R. (2013). Polygenic risk and the developmental progression to heavy, persistent smoking and nicotine dependence: Evidence from a 4-decade longitudinal study. *JAMA Psychiatry, 70*(5), 534–542.

Bonnie, R. J., Stratton, K., & Kwan, L. Y. (Eds.). (2015). *Public health implications of raising the minimum age of legal access to tobacco products*. Washington, DC: National Academies Press. Retrieved December 15, 2020, from http://tobacco.cleartheair.org.hk/wp-content/uploads/2015/08/18997-2.pdf

Brook, J. S., Balka, E. B., Ning, Y., & Brook, D. W. (2007). Trajectories of cigarette smoking among African Americans and Puerto Ricans from adolescence to young adulthood: Associations with dependence on alcohol and illegal drugs. *American Journal on Addictions, 16*(3), 195–201.

Brown, T., Platt, S., & Amos, A. (2014). Equity impact of population-level interventions and policies to reduce smoking in adults: A systematic review. *Drug and Alcohol Dependence, 138*, 7–16.

Cantrell, J., Bennett, M., Mowery, P., Xiao, H., Rath, J., Hair, E., & Vallone, D. (2018). Patterns in first and daily cigarette initiation among youth and young adults from 2002 to 2015. *PloS one*, *13*(8), e0200827.

Chassin, L., Presson, C. C., Sherman, S. J., & Edwards, D. A. (1990). The natural history of cigarette smoking: Predicting young-adult smoking outcomes from adolescent smoking patterns. *Health Psychology, 9*(6), 701.

Chassin, L., Presson, C., Seo, D. C., Sherman, S. J., Macy, J., Wirth, R. J., & Curran, P. (2008). Multiple trajectories of cigarette smoking and the intergenerational transmission of smoking: A multigenerational, longitudinal study of a Midwestern community sample. *Health Psychology, 27*(6), 819.

Chen, P., & Chantala, K. (2014). Guidelines for analyzing Add Health data. *Carolina Population Center, University of North Carolina at Chapel Hill*, *710*. Retrieved on December 16, 2020, from https://addhealth.cpc.unc.edu/wp-content/uploads/docs/user_guides/GuidelinesforAnalysisofAddHealthData_202004.pdf

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.

Fagan, P., Moolchan, E. T., Lawrence, D., Fernander, A., & Ponder, P. K. (2007). Identifying health disparities across the tobacco continuum. *Addiction, 102*, 5–29.

Garrett, B. E., Dube, S. R., Babb, S., & McAfee, T. (2014). Addressing the social determinants of health to reduce tobacco-related disparities. *Nicotine & Tobacco Research, 17*(8), 892–897.

Goldade, K., Choi, K., Bernat, D. H., Klein, E. G., Okuyemi, K. S., & Forster, J. (2012). Multilevel predictors of smoking initiation among adolescents: Findings from the Minnesota Adolescent Community Cohort (MACC) study. *Preventive Medicine, 54*(3–4), 242–246.

Griesler, P. C., Kandel, D. B., & Davies, M. (2002). Ethnic differences in predictors of initiation and persistence of adolescent cigarette smoking in the National Longitudinal Survey of Youth. *Nicotine & Tobacco Research, 4*(1), 79–93.

Harris, K. M., Halpern, C. T., Whitsel, E., Hussey, J., Tabor, J., Entzel, P., & Udry, J. R. (2009). The national longitudinal study of adolescent to adult health: research design. Retrieved on December 16, 2020, from https://addhealth.cpc.unc.edu/documentation/study-design/

Hu, T., Gall, S. L., Widome, R., Bazzano, L. A., Burns, T. L., Daniels, S. R., ... & Prineas, R. J. (2020). Childhood/adolescent smoking and adult smoking and cessation: The International Childhood Cardiovascular Cohort (i3C) Consortium. *Journal of the American Heart Association*, *9*(7), e014381.

Jamal, A., Gentzke, A., Hu, S. S., Cullen, K. A., Apelberg, B. J., Homa, D. M., & King, B. A. (2017). Tobacco Use Among Middle and High School Students - United States, 2011–2016. *MMWR. Morbidity and mortality weekly report*, *66*(23), 597–603. https://doi.org/10.15585/mmwr.mm6623a1

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E. (2019). *Monitoring the future national survey results on drug use, 1975–2018: Overview*. Institute for Social Research.

Kandel, D. B., Kiros, G. E., Schaffran, C., & Hu, M. C. (2004). Racial/ethnic differences in cigarette smoking initiation and progression to daily smoking: A multilevel analysis. *American Journal of Public Health, 94*(1), 128–135.

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology, 72*(6), 558.

King, J. L., Reboussin, D., Ross, J. C., Wiseman, K. D., Wagoner, K. G., & Sutfin, E. L. (2018). Polytobacco use among a nationally representative sample of adolescent and young adult e-cigarette users. *Journal of Adolescent Health, 63*(4), 407–412.

Liu, S., Rovine, M. J., & Molenaar, P. (2012). Selecting a linear mixed model for longitudinal data: repeated measures analysis of variance, covariance pattern model, and growth curve approaches. *Psychological Methods*, *17*(1), 15.

Maruyama, N., Takahashi, F., & Takeuchi, M. (2009). Prediction of an outcome using trajectories estimated from a linear mixed model. *Journal of Biopharmaceutical Statistics, 19*(5), 779–790.

O'Loughlin, J. L., Dugas, E. N., O'Loughlin, E. K., Karp, I., & Sylvestre, M. P. (2014). Incidence and determinants of cigarette smoking initiation in young adults. *Journal of Adolescent Health, 54*(1), 26–32.

Piasecki, T. M., Hedeker, D., Dierker, L. C., & Mermelstein, R. J. (2016). Progression of nicotine dependence, mood level, and mood variability in adolescent smokers. *Psychology of Addictive Behaviors : Journal of the Society of Psychologists in Addictive Behaviors, 30*(4), 484–493. https://doi.org/10.1037/adb0000165

Saddleson, M. L., Kozlowski, L. T., Giovino, G. A., Homish, G. G., Mahoney, M. C., & Goniewicz, M. L. (2016). Assessing 30-day quantity-frequency of US adolescent cigarette smoking as a predictor of adult smoking 14 years later. *Drug and Alcohol Dependence, 162*, 92–98.

Sargent, J. D., Gabrielli, J., Budney, A., Soneji, S., & Wills, T. A. (2017). Adolescent smoking experimentation as a predictor of daily cigarette smoking. *Drug and Alcohol Dependence, 175*, 55–59.

SAS, I. (2013). Base SAS 9.4 procedures guide: statistical procedures. *Cary, NC, USA: SAS Institute Inc*.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*(4), 323–355.

Sung, H. Y., Wang, Y., Yao, T., Lightwood, J., & Max, W. (2018). Polytobacco use and nicotine dependence symptoms among US adults, 2012–2014. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*, *20*(suppl_1), S88–S98. https://doi.org/10.1093/ntr/nty050

Sutter, M. E., Everhart, R. S., Miadich, S., Rudy, A. K., Nasim, A., & Cobb, C. O. (2018). Patterns and profiles of adolescent tobacco users: results from the Virginia youth survey. *Nicotine and Tobacco Research*, *20*(suppl_1), S39-S47.

Tanner, T., Päkkilä, J., Karjalainen, K., Kämppi, A., Järvelin, M. R., Patinen, P., ... & Anttonen, V. (2015). Smoking, alcohol use, socioeconomic background and oral health among young Finnish adults. *Community Dentistry and Oral Epidemiology, 43*(5), 406–414.

Thompson, A. B., Tebes, J. K., & McKee, S. A. (2015). Gender differences in age of smoking initiation and its association with health. *Addiction Research & Theory, 23*(5), 413–420.

Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology, 39*(2), 81–90.

Trinidad, D. R., Gilpin, E. A., Lee, L., & Pierce, J. P. (2004). Do the majority of Asian-American and African-American smokers start as adults? *American Journal of Preventive Medicine, 26*(2), 156–158.

US Department of Health and Human Services. (2012). Preventing tobacco use among youth and young adults: a report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

US Department of Health and Human Services. (2014). The health consequences of smoking—50 years of progress: a report of the Surgeon General. U.S. Dept. of Health and Human Services, Public Health Service, Office of the Surgeon General, Rockville, MD. Retrieved on October 26, 2020, from https://aahb.org/Resources/Pictures/Meetings/2014-Charleston/PPT%20Presentations/Sunday%20Welcome/Abrams.AAHB.3.13.v1.o.pdf

Yoo, W., Mayberry, R., Bae, S., Singh, K., Peter He, Q., & Lillard, J. W., Jr. (2014). A study of effects of multicollinearity in the multivariable analysis. *International Journal of Applied Science and Technology, 4*(5), 9–19.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

David Adzrago[1] · Lucy Kavi[2] · Rosemary I. Ezeugoh[2] · Benedicta Osafo-Darko[3]

[1] Center for Health Promotion and Prevention Research, The University of Texas School of Public Health, Health Science Center at Houston, 7000 Fannin, Suite 2502A, Houston, TX 77030, USA

[2] Maryland Institute of Applied Environmental Health, School of Public Health, University of Maryland, College Park, MD, USA

[3] Department of Behavioral and Community Health, School of Public Health, University of Maryland, College Park, MD, USA