

Smoking Trajectories, Health, and Mortality across the  
Adult Lifespan

by

Zachary Adam Koel Frosch  
Class of 2008

A thesis submitted to the  
faculty of Wesleyan University  
in partial fulfillment of the requirements for the  
Degree of Bachelor of Arts  
with Departmental Honors in Psychology

### **Acknowledgements:**

I would like to thank the many people who made this project possible. My adviser, Professor Lisa Dierker, guided me through every step of the study, both in conducting the analysis and crafting the paper. Without Dr. Robert Waldinger, who made the study data available, this project would not have been possible. Professor Jen Rose helped to both conduct the trajectory analysis and interpret its results. Sally Kuo and Melissa Pergakis helped me to interpret the surveys and in retrieving the biographical and medical information of the participants. I would also like to thank Hannah Nam for providing her thoughts on the paper. Finally, thank you to my friends and family for their support and for listening to me talk about this project probably more than they wanted to these last two years.

### **Abstract:**

Research has clearly established that smoking negatively impacts one's health. However, previous work examining smoking trajectories has neither followed participants past early adulthood nor studied how these trajectories may link to specific health outcomes. The present study aims to address these shortcomings by using latent class growth modeling to examine smoking trajectories and health from early adulthood through late life in a cohort of 232 Harvard graduates with a mean birth year of 1921. The present analysis showed five distinct smoking trajectories based largely on the age that participants quit smoking. The trajectories were significantly associated with both chronic lung diseases and mortality and demonstrated that the later a participant quit smoking, the higher the probability of developing lung disease, and of premature death.

## **Introduction:**

As the single most preventable cause of premature death in the United States, smoking results in 440,000 deaths each year, translating to 5.6 million years of potential life lost and 75 billion dollars in direct health care costs (MMWR, 2003). Numerous retrospective reports indicate that smoking is associated with specific negative health outcomes, including several types of cancer as well as cardiovascular and respiratory diseases such as coronary heart disease (CHD), stroke, aneurysm, atherosclerosis, chronic obstructive pulmonary disease (COPD) and pneumonia (USDHHS, 2004). Prospective work has confirmed many of these associations (Engeland, Haldorsen, Anderson, & Tretli, 1996; Freund, Belanger, D'Agostino, & Kannel, 1993; Giovannucci, 1994; Godtfredsen, Prescott, & Osler, 2005; Hirdes, Brown, Vigoda, Forbes, & Crawford, 1987; Howard et al., 1998; Lam et al., 1997; Nusselder, Looman, Mheen, Mheen, & Mackenbach, 2000; Ostbye & Taylor, 2004; Ostbye, Taylor, & Jung, 2002; Prescott, Hippe, Schnohr, Hein, & Vestbo, 1998; Prescott et al., 1998; Simons, Simons, McCallum, & Friedlander, 2005; Wannamethee, Shaper, & Perry, 2001; Weir, 1970; Yuan et al., 1996) and has demonstrated a temporal ordering that suggests the likelihood of a causal relationship between smoking and these disease outcomes.

Unfortunately, such prospective research is limited in that it typically has been conducted over short periods of time, with few studies administering repeated smoking assessments for longer than a decade (Giovannucci, 1994; Howard et al., 1998; Ostbye, Taylor, & Jung, 2002; Weir, 1970; Yuan et al., 1996). For example, the

Health and Retirement Survey (HRS) and the Asset and Health Dynamics Among the Oldest Old (AHEAD), followed participants for an average of seven to eight years (Ostbye & Taylor, 2004; Ostbye, Taylor, & Jung, 2002). Among studies with longer periods of follow up, smoking is often only assessed at baseline (Engeland, Haldorsen, Anderson, & Tretli, 1996; Enstrom, 1999; Hirdes, Brown, Vigoda, Forbes, & Crawford, 1987; Lam et al., 1997; Rogot & Murray, 1989). For example, the U.S. Veterans Study followed a group of over 200,000 veterans from 1954 through 1979. Even though data on mortality was obtained at each longitudinal assessment, data on smoking behavior was obtained only at the time participants entered the study (Enstrom, 1999; Rogot & Murray, 1989). Such studies lack a complete picture of participants' smoking history. Further, among those studies of United States populations with the longest periods of repeated longitudinal smoking assessment (i.e. the Framingham Study (Freund, Belanger, D'Agostino, & Kannel, 1993; Freund, 1992; Gordon, 1975) and the First Cancer Prevention Study, CPS-I (Knoke, Shanks, Vaugh, Thun, & Burns, 2004; Thun, Day-Lally, Calle, Flanders, & Heath, 1995; Thun & Heath, 1997)), baseline smoking behavior is most often captured upon study entry through retrospective reports by adults who have long passed the age of risk for smoking initiation and established use. Further, participants are enrolled at different ages and are followed for different periods of their lives (Freund, Belanger, D'Agostino, & Kannel, 1993). As a result, their risk for chronic illness during the observation period is not comparable across studies. For example, in the Framingham study, participants entered between age 30 and 62 and, though information on smoking behavior during the follow-up was obtained biennially, baseline smoking

was obtained retrospectively (Freund, Belanger, D'Agostino, & Kannel, 1993). Further, despite the repeated smoking assessments administered in some longitudinal studies, statistical analysis are most often based either on current use (Freund, Belanger, D'Agostino, & Kannel, 1993) or on baseline use (Knoke, Shanks, Vaugh, Thun, & Burns, 2004; Thun, Day-Lally, Calle, Flanders, & Heath, 1995; Thun & Heath, 1997). Because most smokers quit and relapse several times before abstaining permanently, short term and compressed measurements of smoking are necessarily simplifications of the quitting process (USDHHS, 1990), and thus fail to maximally characterize changes in smoking history across the lifespan or to determine how such changes may uniquely impact the development of chronic disease.

One approach that addresses such limitations is the latent class growth model (LCGM). This technique is a type of finite mixture model developed to identify population subgroups following distinct behavioral trajectories (Methuén & Methuén, 2000; Nagin, 1999). This method identifies a number of typical trajectories for the examined variable. While it is understood that not all individual trajectories will perfectly conform to group trajectories, analysis of these group trajectories has proved meaningful in both a clinical and a research setting (Nagin, 1999). LCGM is semi-parametric and group-based. It uses multinomial modeling in order to identify unique groups of individual patterns of behavior (Delucchi, Matzger, & Weisner, 2004).

While prospective work based on this type of analysis has previously defined trajectories of smoking behavior within adolescence and across early adulthood, as well as examined precipitating risk factors that increase the probability of following a particular trajectory, this research has not examined trajectories across the lifespan,

nor has it linked smoking trajectories with disease outcomes. To address these shortcomings, the present study aims to extend research on the association between smoking behavior and chronic disease in the following ways: (1) by following a cohort from the time of initiation of regular smoking patterns (i.e., age 21) into old age (i.e., age 82), when chronic illness is most likely to impact health and longevity; and (2) by incorporating chronic disease outcomes into the trajectory analysis to evaluate whether an individual's tendency to follow a particular smoking trajectory is associated with specific chronic diseases and longevity.

## **Methods**

The present analysis utilized a sample known as the Grant Study, selected from the Harvard classes of 1942-1944 (Vaillant, 1979). Participants had a mean year of birth of 1921 (SD=2 years) and were selected for the study as sophomores from 1938-1942 (Vaillant, 1996). The final 267 participants were selected on the basis of freshman health examinations which indicated no mental or physical problems, a freshman grade point average above C, and judgment by the dean's office and dormitory staff that they were free of physical, emotional and academic difficulties (Vaillant, Schnurr, Baron, & Gerber, 1991; Vaillant, 1979, 1996, 1998). The sample was all male and exclusively white, with the exception of one Hispanic participant.

Participants were followed by annual or biennial questionnaires from age 21 (Vaillant, 2003) through 2003 (age  $82 \pm 2$  years). Questionnaires that inquired about smoking behavior were administered annually (1940 through 1947 and 1952 through

1955), biennially (1947 through 1951, 1970 through 1972, 1975 through 1977 and 1987 through 2003), or every three years (1957 through 1960, 1967 through 1970 and 1972 through 1975). For this study, annual data from 1940-1945 and data from the years 1999, 2001, and 2003 were excluded due to more than 50% missing data, leaving a total of 22 time points between 1946 and 1997. Eighteen men who were missing tobacco data at all 22 time points were also excluded, as were an additional 17 men who were missing tobacco data on 17 or more of the 22 time points (95th percentile or higher for amount of missing data). Thus, the analytic sample included 232 men, of whom 18.5% had no missing tobacco data. The median number of time points missing tobacco data in this sample was 4.5 (range 0-16).

Questionnaires included items inquiring whether or not the individual smoked cigarettes, pipes, or cigars and, if so, how many were smoked per day. Participants were able to indicate a number of discrete categories for each type of tobacco used, ranging from no use to six or more pipes or cigars per day and no use of cigarettes to smoking more than two packs per day. An aggregate tobacco variable was created for each year by combining the three types of tobacco use. Participants were considered daily smokers for that year if they indicated that they had smoked any type of tobacco on a daily basis. Non-smokers included those participants who indicated no use of any type of tobacco during that year and those indicating that their use of tobacco products was less than daily. Only five participants reported any weekly or more frequent tobacco use after lifetime cessation of daily use, and these participants reported such use on four or fewer data points. Although response categories varied slightly during longitudinal assessment (Appendix 1), these differences did not affect



the aggregate tobacco variables (i.e. measurement of the presence or absence of daily tobacco exposure).

For the present analysis, the physical health of the men was coded for a number of disease categories. An internist reviewed each participant's physical examinations, which were obtained every five years, and rated each individual from one (excellent health) to five (deceased). A rating of two denoted minor chronic problems, three denoted serious chronic illness without disability and four denoted irreversible illness with serious disability (Vaillant, 1998). The internist also noted specific diagnoses that the participant had developed at the time of each rating. The internist's ratings and diagnoses for the most recent physical examination were reviewed and presence of the following health problems was considered: cancer, heart disease, vascular disease, cholesterol disease, lung disease, diabetes, cerebrovascular disease, and hypertension. Diagnoses that resulted in a participant being coded as positive for a particular disease category included: *cancer*: cancer, CA, carcinoma, adenoma, malignant lymphoma, melanoma, chronic lymphocytic leukemia, basal cell epithelioma, glioma; *heart disease*: myocardial infarction, coronary artery disease, congestive heart failure, cardiomyopathy, graft or angioplasty of coronary artery, angina, ischemia; *vascular disease*: peripheral vascular disease, aortic aneurysm, angioplasty of any peripheral vessel, deep vein thrombosis, carotid or aortic stenosis, varicosities, carotid inclusion, carotid artery surgery, aortic insufficiency, aortic sclerosis, Ca<sup>++</sup>, carotid endarectomy, bypass of a peripheral vessel, lateral ischemia, vascular dementia, vessel stent, venous stasis, pulmonary embolism; *cholesterol disease*: hypercholesterolemia, high lipids, high HDL, low HDL, high LDL,

dyslipidemia, high cholesterol, hyperlipidemia; *diabetes*: diabetes mellitus, glucose intolerance, glucose tolerance, high glucose; *cerebrovascular disease*: cerebrovascular accident, transient ischemic attack, infarct, sub-arachnoid hemorrhage, brain hemorrhage; *hypertension*: hypertension, high blood pressure, on antihypertensives; and *lung disease*: chronic obstructive pulmonary disease, hyperinflated lungs, pneumonia, tuberculosis, pulmonary edema, emphysema, lung resection, lung abscess. If any of these diagnoses were listed as a cause of death on a participant's death certificate, the condition was also coded as present. If two readers deemed a diagnosis unreadable, one of who was a medical doctor blind to other study data, the diagnosis was not coded.

Descriptive statistics were obtained for a number of smoking variables in order to describe the smoking behavior of the sample as a whole. Subsequently, participants were grouped into different trajectories based on a latent class growth curve analysis (LCGA). LCGA seeks to identify groups or classes of individual who are similar to each other in their smoking behavior over time, that is, to identify groups of individuals who share similar smoking patterns (Methuén & Methuén, 2000; Nagin, 1999). Variability in smoking behavior over time is modeled by a latent categorical variable whose categories reflect different latent trajectory classes. For this study, age at each time of measurement was used to formulate age categories. Such formulation ensured that an adequate number of participants were represented in each category.

LCGA proceeds by testing a series of models specifying a different number of latent classes. In this study, models specifying two through six classes were tested.

Choosing the model with the number of classes that best represents the true number of classes is guided by expectations based on previous research, assessment of model fit using statistical criteria, parsimony, and additional analyses to determine if the trajectory classes are distinct and meaningful. Several recommended criteria were used to assess model fit (Nagin, 1999, 2005). The Bayesian Information Criteria (BIC) is a standard method for deciding whether additional classes result in a better fitting model, where models with BIC values closest to zero are favored. The Lo-Mendell-Rubin likelihood ratio provides a test that compares the current model to a model with one less class where a p-value of less than 0.05 suggests rejecting the null hypothesis of retaining a model with one less class (Lo, Mendell, & Rubin, 2001). Finally, we examined the model estimated (posterior) probabilities of group membership, which is an average of each individual's likelihood of belonging to each class. Ideally, individuals' posterior probabilities equal one for their most likely trajectory group and zero for the remaining trajectory groups, however a minimum average posterior probability of .70 for all trajectory groups is considered evidence of acceptable model fit (Nagin, 2005).

After deciding which model to retain, chi-square analyses were run to compare a number of discrete descriptive smoking characteristics and disease outcome variables between the trajectory groups. Where a significant result was found, post-hoc paired comparisons were conducted. For continuous variables, a one-way ANOVA was used to compare group means followed by post-hoc paired comparisons using Tukey's HSD.

Finally, a discrete time survival LCGA was conducted to evaluate whether progression in the probability of mortality (mortality hazard rate) varied across latent trajectory classes. This analysis simply extended the LCGA model to incorporate a variable that indicated whether or not participants had died at each time point. Participants who died or dropped out of the study were censored at subsequent time points. The survival part of the LCGA provides a test for significant differences in mortality hazard rates between trajectory classes.

## **Results**

A total of 25% (n=59) of the sample reported no daily smoking at any of the assessment waves. Among those with exposure to daily smoking, 36% (n=62) smoked only one type of tobacco product during their lifetime (28% (n=49) cigarettes only, 5% (n=9) pipes only and 2% (n=4) cigars only). The remaining 64% of daily smokers (n=111) reported smoking multiple types of tobacco during their lifetime (26% (n=45) cigarettes and pipes, 8% (n=14) cigarettes and cigars, 11% (n=19) pipes and cigars and 19% (n=33) at least some daily use of all three tobacco products). Among those who reported daily use of more than one tobacco product, this typically represented changes from one product to another over time, rather than simultaneous use of multiple products. That is, while simultaneous use of multiple tobacco products was reported on one or more assessment occasions for 80% (n=88) of those individuals who reported smoking more than one type of tobacco in their lifetime, only 13% (n=14) of multiple tobacco users did so on a majority of daily smoking occasions.

A total of 80% of smokers and 59% of the overall sample (n=137) had begun smoking by the first assessment in 1946. Another 5% of smokers had started by 1947, and by the end of the 1950s, 96% of those who would smoke daily at any point during their lifetime had initiated daily use. The median age when participants started daily smoking was 26, and 85% of daily smokers had initiated the habit by the end of their 20s. Among all daily smokers in the sample (n=173), almost a quarter had quit by 1964 (~ age 40), half by 1977 (~ age 60), and 80% by the last year of observation (~ age 80).

Among daily smokers, 41% (n=71) quit smoking<sup>1</sup> during their lifetime and did not relapse at any subsequent assessment; 39% (n=68) quit smoking, but reported at least one relapse during subsequent assessments, and the remaining 20% (n=34) reported chronic exposure, where following onset, daily use of tobacco was reported at each assessment preceding the participant's death. The number of times a relapsing smoker quit daily smoking ranged from one to five with a mean of 2.28 times. Overall, smokers quit daily smoking a mean 1.31 times.

Based on model fit statistics and parsimony, a five trajectory class model was chosen to represent smoking across the lifespan (Table 1). Although the adjusted Lo-Mendell-Rubin test suggested retaining a six-class model over a five-class model, there was virtually no improvement in the BIC for the six-class model, and the average posterior probabilities dropped below 0.70 for two of the classes, which suggested a loss in classification quality in the six-class model. The average latent

---

<sup>1</sup> Quitting daily smoking was defined as a data point where a participant reported no daily use of tobacco following one or more assessments where the participant reported daily smoking behavior.

trajectories for the five distinct groups of smokers are shown in Figure 1. There were a small number of “lifetime users” (n=11; average posterior probability = 0.94) who maintained tobacco use across their lifespan. “Late quitters” (n=48; average posterior probability = 0.71) smoked for a large portion of their lives and quit at a median age of 69<sup>2</sup>. Tobacco use for “middle quitters” (n=69; average posterior probability = 0.85) declined steadily over the lifetime and they quit at a median age of 56, while “early quitters” (n=26; average posterior probability = 0.70) ceased daily smoking at a much earlier age than other groups (i.e. median age of 40). There was also a group of “low users” (n=78; average posterior probability = 0.95) whose tobacco use remained minimal across measurements. As a group, low users reported daily smoking at only 2% of time points. The validity of the trajectory classes was demonstrated by an ANOVA comparing the year that smokers quit by trajectory class ( $F(4, 134)=35.4$ ,  $p=.0001$ ) which showed that late quitters quit later than middle and early quitters, and that middle quitters quit later than early quitters. Table 2 shows descriptive variables for each of the trajectory classes.

Table 2 also summarizes the disease outcomes for the different trajectory classes. Differences between groups were observed only for lung disease. Post-hoc paired comparisons showed that late quitters had a higher prevalence of lung disease than low users and early quitters and that middle quitters experienced a higher prevalence than early quitters. No differences in rates of the remaining disease categories were found between the trajectory classes.

---

<sup>2</sup> All medians are for those who quit during their lifetime.

Results of the discrete time survival latent class growth analysis are presented in Table 3 and Figure 2. The hazard rate for death was significantly higher for middle and late quitters compared to low users. The hazard rate for death was significantly lower for early quitters compared to middle and late quitters. Survival curves indicate that low users had a higher survival probability compared to the three quit groups by age 81-85, though when compared with early quitters the difference was not significant. Late quitters were more likely to die during the study and to die earlier, with the lowest survival probability by age 65-69 compared to the other classes and with significantly higher hazard rates than early quitters and low users. Survival curves were similar for early and middle quitters, but by age 81 to 85, middle quitters had lower survival probability and a significantly higher hazard rate for death. Because of the small number of “lifetime users” (n=11), it is difficult to draw any conclusions relating to this group. As a result, this group is not discussed here and further investigation with a larger number of participants is suggested.

## **Discussion**

This study extends research on the association between smoking behavior and chronic disease by following a male cohort from the time of initiation of regular smoking behavior into old age, when chronic illness is most likely to impact health and longevity. Because this sample includes assessment of smoking at several points over the participants’ lifetime, it presents a unique opportunity to examine the natural history of smoking behavior as it links to longevity and disease. Three major findings emerged. First, the sample contained five distinct smoking trajectories based on the

age at which participants quit daily smoking. Second, those participants following smoking trajectories with later cessation had a higher likelihood of developing lung disease than those with who quit smoking at an earlier age. Finally, participants following trajectories with later cessation lived shorter lives.

The vast majority of the sample spent at least some of their life smoking daily. Seventy-five percent of this sample smoked daily at some point during the observation period, which is higher than the 50% lifetime prevalence of smoking<sup>3</sup> in the current United States male population (Adams & Schoenborn, 2006). However, 62% of the current population of males over the age of 75 have ever smoked (Adams & Schoenborn, 2006). Thus, though the prevalence of smoking in the general US population is still less than the 75% of men who smoked daily in the current sample, because of the higher mortality rate among smokers, it would be expected that fewer smokers than non-smokers would be alive at such an advanced age, reducing the apparent prevalence of smoking in the overall population.

Research has shown that less than a fifth of current smokers begin smoking after age 21 (Adams & Schoenborn, 2006). This early initiation of smoking habits is consistent with the current sample, in which most smokers started before age 30. The majority of the smokers in the sample quit daily smoking at least once during the observation period. Half had quit by their mid 50s, and 80% by the end of their 60s. Though research has shown that 40% of daily smokers attempt to quit each year (Adams & Schoenborn, 2006), quitting smoking is a difficult process and the majority of smokers relapse following any particular cessation attempt (USDHHS, 1990).

---

<sup>3</sup> Defined as smoking a total 100 or more cigarettes over the life course.



Though relapse is much less common after two years of abstinence, 10% of smokers who have quit for two years or more suffer a relapse (USDHHS, 1990). While other studies have shown that the average smoker attempts to quit three or four times before doing so successfully (USDHHS, 1990), because most smoking assessments in this study were biennial, the present analysis is only able to assess long term quitting. This could explain why participants in the present study quit with a mean much lower than that of the average smoker and why about the same number of smokers quit without relapsing as those who quit and relapsed.

While other trajectory work has followed participants through adolescence and early adulthood (Chassin, Presson, Pitts, & Sherman, 2000), to our knowledge this is the first to examine trajectories of smoking across the adult lifespan. These trajectory analyses based on LCGM identified five smoking trajectories based largely on the timing of lifetime cessation. Low users were at reduced risk for lung diseases compared to late quitters, had a lower probability of death than middle and late quitters, and were more likely to live past age 80 compared to each of the quit groups, though the difference between early quitters and low users was not significant. The early quitters stopped smoking by a median age of 40 and had a lower prevalence of lung disease across the life course compared to middle and late quitters. Furthermore, early quitters lived longer than both middle and late quitters, who had median quitting ages of 56 and 69 respectively. Late quitters lived significantly shorter lives than both middle quitters and low users, having a higher probability of death before age 70. Comparing the survival curves of middle and late quitters (Figure 2) shows that, by age 81 to 85, late quitters were more likely to have died than middle quitters. Though

this difference was not significant in the present sample, it is possible that with a larger sample size a significant difference would be observed. Although the survival curve for lifetime smokers suggests a higher probability of survival for this group, this result must be viewed with caution due to the small number of participants in this class. It is likely that the estimated survival curve based on this small sample is not representative of the true lifetime survival probability for the population of lifetime smokers in general.

Overall, these findings generally confirm previous research documenting the health benefits of smoking cessation. For example, ex-smokers have been shown to have a reduced risk of mortality associated with cancers, cardiovascular diseases, and respiratory diseases (MMWR, 1993). Large studies show that smokers who quit 15 years previously are not more likely than those who never smoked to be in poor health (Ostbye, Taylor, & Jung, 2002) and that 10 to 15 years after quitting the risk of mortality returns to nearly the level of nonsmokers (USDHHS, 1990). Previous work has also shown the benefits of quitting even later in life. Those who quit at age 50 are half as likely to die within 15 years than those who continue to smoke (USDHHS, 1990). The present study confirms the benefits of quitting and demonstrates that there is a protective effect of early smoking cessation. After many years of abstinence, the risk of mortality for smokers may return to that of non-smokers, as the survival curve analysis shows that early quitters were not significantly more likely to die earlier than low users. However, early quitters had a significantly higher survival probability than middle quitters at age 81-85, over 25 and 40 years after the average age of quitting for these groups, contrasting with previous studies that have shown that risk of mortality

is similar in those who quit smoking 15 years previously. This result implies that the earlier a smoker quits, the greater the health benefits, and that these benefits are observed even decades after smoking cessation.

This study is based on a small sample of 232 participants. A larger sample may have provided the power necessary to replicate many of the previously demonstrated associations. All of the diseases examined in the present study that did not show a significant association with smoking trajectories have been shown previously to be associated with smoking behavior in other populations. However, such diseases can be misdiagnosed and thus the potential for measurement error is much higher than when measuring mortality. Thus, despite the non-significant results for many of the chronic disease categories, the association between smoking trajectory and mortality in this sample indicates that trajectory analysis of smoking patterns is a good predictor of health. Because of the nature of the sample, further investigation is needed to determine whether the current findings can be replicated in other cohorts, and whether some of the results that approach significance in this study would prove significant with a larger sample. However, the long-term nature of the study, the recurrent smoking and health assessments, and the low dropout rates make the study a difficult one to replicate.

Many of the other limitations of this study derive from the homogeneity of the sample and have been outlined previously (Vaillant, 1979, 1998). The sample is a single small, educationally and racially homogeneous cohort that has been pre-selected for good health. This suggests that these men are likely to remain healthy longer than the population at large. The median life expectancy for white men born in

1920 is 55 years (Hill, 1936) compared to the average lifespan of 74 in the present sample among the 45% of participants who already died. However, the selection of these men for their health minimizes the possibility of a pre-existing condition and thus maximizes the probability that the conditions examined were developed during the observation period. Moreover, while the homogeneity of the sample limits the generalizability of the results, it eliminates the need to control for factors such as education, socio-economic class, race, and gender.

The sample has many other advantages. The length of the observation and the repeated assessment of smoking behavior, including the type of tobacco used, allows for a more complete picture of the participants' smoking histories. Additionally, as evidenced by the 39% of participants who quit smoking and then relapsed at least once, quitting smoking is a dynamic process (USDHHS, 1990) and an analysis such as LCGM that examines the long-term smoking habits of participants is more likely to accurately characterize quitting behavior. Finally, the length of the follow-up allows for the examination of the outcomes of smoking behavior, and for the inclusion of data from late life, when chronic diseases such as those examined are most likely to have developed.

Unlike previous longitudinal research, including CPS-I and the Framingham Study, which focused exclusively on cigarette smoking (Freund, Belanger, D'Agostino, & Kannel, 1993; Gordon, 1975; Knoke, Shanks, Vaughn, Thun, & Burns, 2004), the present study examined tobacco exposure through pipes and cigars in addition to cigarettes. Two-thirds of smokers were found to use multiple types of tobacco, though typically in a serial manner, as opposed to concurrently. Despite the

high prevalence of exposure to tobacco in multiple forms, most smokers smoked at least some cigarettes, with the two highest combinations of use being cigarettes only (28% of smokers, n=49) and cigarettes and pipes (26% of smokers, n=45). However, while cigarette smoking was the dominant form of tobacco use, it does not give a complete picture of a participant's exposure to smoking. After cigarettes only and cigarettes and pipes, the tobacco use group with the most participants was pipes and cigars (19% of smokers, n=33). Thus ignoring the potential impact of pipe and cigar smoking could result in misclassification of some pipe and cigar smokers. Aside from the continued availability of pipes and cigars, new methods of tobacco delivery are now in common usage. Hookahs (or water pipes) have emerged as the newest tobacco trend (ALA, 2007). Despite the exposure to the highly addictive nicotine that occurs while smoking hookah, 90% of hookah smokers believe that it is less addictive than smoking cigarettes (ALA, 2007). Evidence shows that hookah smoking is just as harmful as cigarette smoking and has been linked to many of the same diseases as smoking cigarettes (ALA, 2007). Because of the continued prevalence of multiple methods of tobacco administration, it remains important to examine tobacco exposure beyond cigarette use.

In addition, though smoking prevalence in the United States has declined significantly over the lifetime of the current cohort (NCHS, 2006), rates of nicotine dependence have increased (Breslau, Johnson, Hiripi, & Kessler, 2001). Because nicotine dependence increases the likelihood of continued smoking behavior (Breslau, Johnson, Hiripi, & Kessler, 2001), the impact of smoking on the health of Americans is likely to remain considerable.

## Tables, Figures, and Appendices

Table 1. Fit indices for 2-6 class LCGA models

# of classes	BIC	# of parameters	Range of average posterior probabilities	P-value for $H_0$ =one less class
2	1508.83	5	0.97 - 0.97	0.0000
3	1443.66	8	0.83 – 0.95	0.0033
4	1424.11	11	0.74 – 0.95	0.0078
5	1424.95	14	0.70 – 0.95	0.0352
6	1424.23	17	0.68 – 0.93	0.0042

Table 2: Smoking characteristics and chronic disease by smoking trajectory n (%)

<b>Smoking Characteristics</b>	Lifetime Users n=11	Late Quitters n=48	Middle Quitters n=69	Early Quitters n=26	Low Users n=78	Statistics
Smoking in 1946	n=9 (82%)	n=44 (96%)	n=55 (83%)	n=19 (83%)	n=9 (82%)	$X^2=116$ , df=4, p=.0001 <sup>4</sup>
Used Multiple Types of Tobacco	n=11 (100%)	n=34 (71%)	n=47 (68%)	n=13 (50%)	n=6 (8%)	$X^2=172$ , df=8, p=.0001 <sup>5</sup>
Used Multiple Types Simultaneously at Least Once	n=11 (100%)	n=28 (58%)	n=35 (51%)	n=10 (39%)	n=4 (5%)	$X^2=66.9$ df=4, p=.0001 <sup>6</sup>
Used Multiple Types Simultaneously on Majority of Points was Smoking	n=2 (18%)	n=3 (6%)	n=5 (7%)	n=1 (4%)	n=3 (4%)	$X^2=3.92$ , df=4, p=.416

<sup>4</sup> Post-Hoc tests demonstrated that fewer low users were smoking daily in 1946 than all other groups and more late quitters were smoking daily in 1946 than middle quitters.

<sup>5</sup> Post-Hoc tests demonstrated that low users smoked multiple types of tobacco during their lifetime daily at a lower rate than all other groups and that lifetime users smoked multiple types at a higher rate than all other groups.

<sup>6</sup> Post-Hoc tests demonstrated that late quitters, middle quitters and early quitters all smoked concurrently at a higher rate than low users and that lifetime users smoked multiple types of tobacco simultaneously at a higher rate than all other groups.

<b>Chronic Disease</b>						
Cancer <sup>7</sup>	n=3 (27%)	n=22 (46%)	n=28 (41%)	n=15 (58%)	n=36 (47%)	X <sup>2</sup> =3.74, df=4, p=.442
Heart Disease <sup>8</sup>	n=2 (18%)	n=19 (40%)	n=24 (35%)	n=6 (23%)	n=18 (23%)	X <sup>2</sup> =5.80, df=4, p=.215
Vascular Disease <sup>9</sup>	n=3 (27%)	n=10 (21%)	n=12 (17%)	n=6 (23%)	n=10 (13%)	X <sup>2</sup> =2.70, df=4, p=.609
Cholesterol Disease <sup>10</sup>	n=5 (46%)	n=12 (25%)	n=18 (26%)	n=9 (34%)	n=21 (27%)	X <sup>2</sup> =2.58, df=4, p=.631
Diabetes <sup>11</sup>	n=2 (18%)	n=6 (13%)	n=6 (9%)	n=3 (12%)	n=4 (5%)	X <sup>2</sup> =3.39, df=4, p=.495
Cerebrovascular Disease <sup>12</sup>	n=0 (0%)	n=8 (17%)	n=6 (9%)	n=3 (12%)	n=12 (15%)	X <sup>2</sup> =3.92, df=4, p=.417
Hypertension <sup>13</sup>	n=5 (46%)	n=19 (40%)	n=22 (32%)	n=12 (46%)	n=32 (42%)	X <sup>2</sup> =2.44, df=4, p=.655
Lung Disease <sup>14</sup>	n=2 (18%)	n=17 (35%)	n=18 (26%)	n=1 (4%)	n=11 (14%)	X <sup>2</sup> =13.7, df=4, p=.008 <sup>15</sup>

<sup>7</sup> cancer, CA, carcinoma, adenoma, malignant lymphoma, melanoma, chronic lymphocytic leukemia, basal cell epithelioma, glioma

<sup>8</sup> myocardial infarction, coronary artery disease, congestive heart failure, cardiomyopathy, graft or angioplasty of coronary artery, angina, ischemia

<sup>9</sup> peripheral vascular disease, aortic aneurysm, angioplasty of any peripheral vessel, deep vein thrombosis, carotid or aortic stenosis, varicosities, carotid inclusion, carotid artery surgery, aortic insufficiency, aortic sclerosis, Ca<sup>++</sup>, carotid endarectomy, bypass of a peripheral vessel, lateral ischemia, vascular dementia, vessel stent, venous stasis, pulmonary embolism

<sup>10</sup> hypercholesterolemia, high lipids, high HDL, low HDL, high LDL dyslipidemia, high cholesterol, hyperlipidemia

<sup>11</sup> diabetes mellitus, glucose intolerance, glucose tolerance, high glucose

<sup>12</sup> cerebrovascular accident, transient ischemic attack, infarct, sub-arachnoid hemorrhage, brain hemorrhage

<sup>13</sup> hypertension, high blood pressure, on antihypertensives

<sup>14</sup> chronic obstructive pulmonary disease, hyperinflated lungs, pneumonia, tuberculosis, pulmonary edema, emphysema, lung resection, lung abscess

<sup>15</sup> Post-hoc tests showed that late quitters experienced more lung disease than low users and than early quitters and that middle quitters experienced more lung disease than early quitters.

Table 3: Latent trajectory class differences in death hazard rate

Class Comparison	Hazard	p
Early Quitters vs. Low Use	-0.07	0.872
Middle Quitters vs. Low Use	1.28	0.001
Late Quitters vs. Low Use	1.09	0.002
Early Quitters vs. Middle Quitters	-1.350	0.011
Early Quitters vs. Late Quitters	-1.158	0.017
Middle Quitters vs. Late Quitters	0.192	0.604

Figure 1. Estimated growth curve trajectories

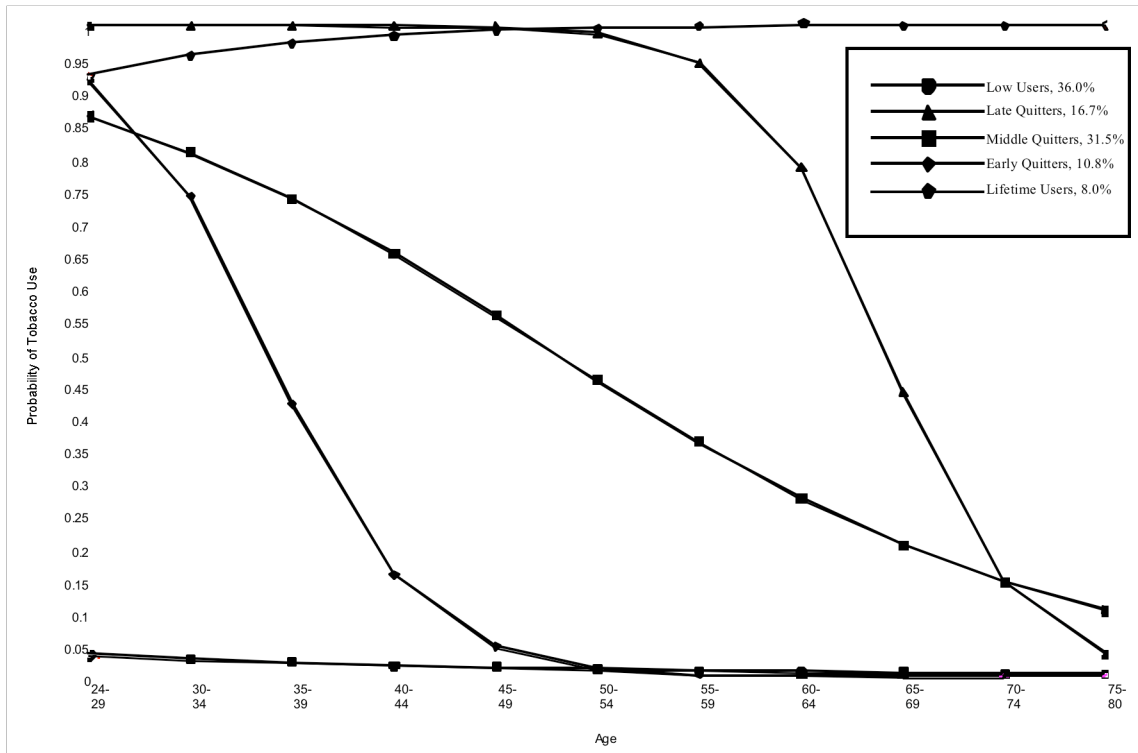
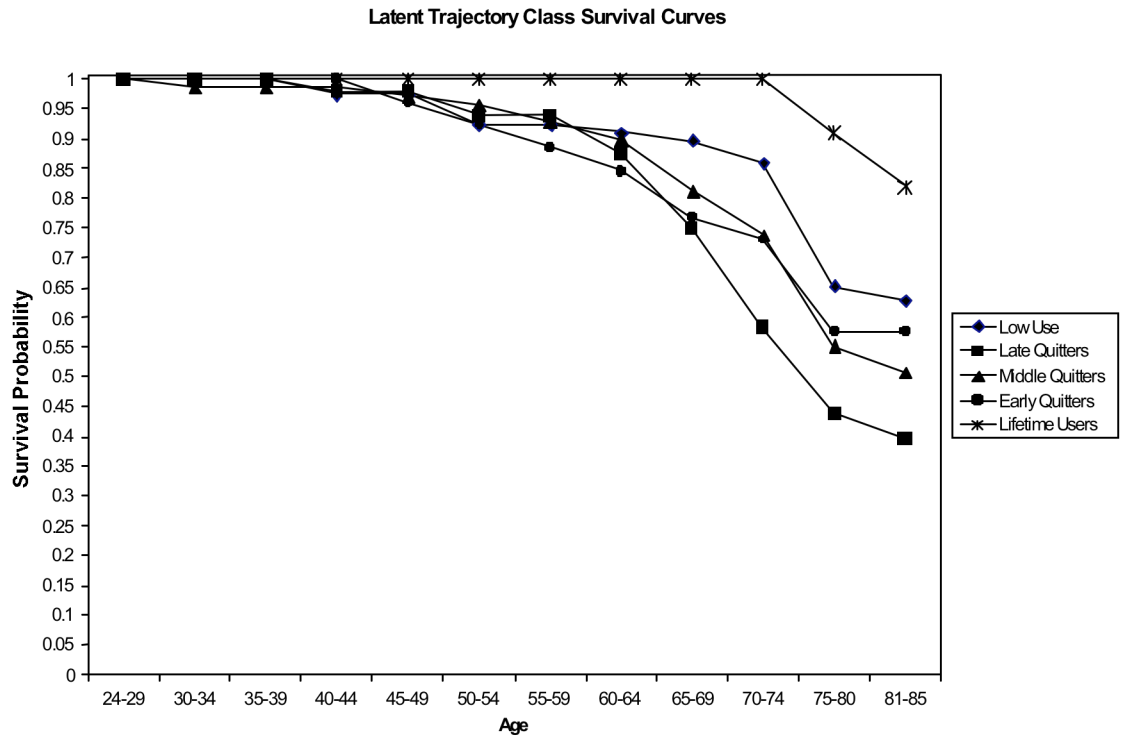




Figure 2: Discrete time latent trajectory class survival curves



Appendix 1: Smoking assessment measures by year

Years	Response Categories for Cigarettes	Response Categories for Pipes	Response Categories for Cigars
1946, 1947, 1949, 1951	none, 0-5, 6-10, 11-19, 1 pack, 2+ packs	0, 1, 2, 3, 4, 5+	0, 1, 2, 3, 4, 5+
1952, 1953, 1954	0-5, 6-10, 11-19, 1 pack, 2+ packs	0, 1, 2, 3, 4, 5, 6	0, 1, 2, 3, 4
1955	0, 1-5, 6-10, 11-19, 1 pack, between 1 and 2 packs, 2+ packs	0, 1, 2, 3, 4, 5, 6	0, 1, 2, 3, 4
1957, 1960	0, 1-5, 6-10, 11-19, 1 pack to 2 packs, 2+ packs	0, 1, 2, 3, 4, 5, 6+	0, 1, 2, 3, 4, 5, 6+
1964, 1967, 1970, 1972, 1975	0, 1-5, 6-10, 11-19, 1 pack, between 1 pack and 2 packs, 2+ packs	0, 1, 2, 3, 4, 5, 6+	0, 1, 2, 3, 4, 5, 6+

1977	0, 1-10, 11-19, 1 pack, between 1 and 2 packs, 2 packs to 3 packs, 3+ packs	0, 1, 2, 3, 4, 5, 6+	0, 1, 2, 3, 4, 5, 6+
1987, 1989, 1991, 1993, 1995, 1997	0, 1-10, 11-19, 1 pack, between 1 and 2 packs, 2 packs to 3 packs, 3+ packs	no cigarettes but I smoke pipes or cigars	no cigarettes but I smoke pipes or cigars

Appendix 2:

Comparison of quitters, relapsers and chronic smokers with trajectory classes:

The lifetime use trajectory was almost totally chronic smokers (81% of the trajectory group, n=9). The late quitters were more evenly split between quitters (50%, n=24) and chronic smokers (38%, n=18). Middle quitters were mostly relapsers (64%, n=44) with quitters being the second highest percentage (28%, n=19). Early quitters were mostly quitters (62%, n=16) and relapsers (39%, n=10). Low users were, for the most part, non-smokers (76%, n=59). Examining to which trajectory the each of the types of smoker belonged to revealed that all non-smokers (100% of the labeled group, n=59) were low users. Quitters were split between late quitters (34% n=24), middle quitters (27%, n=19), early quitters (23%, n=16) and low users (16%, n=11). Relapsers were mostly middle quitters (65%, n=44). Chronic smokers were mostly late quitters (53%, n=18), but also had a notable number of lifetime users (27%, n=9).

## References

- Adams, P., & Schoenborn, C. (2006). Health behaviors of adults: United States, 2002-04. *National Center for Health Statistics. Vital Health Stat, 10*(230).
- ALA. (2007). An Emerging Deadly Trend: Waterpipe Tobacco Use. *American Lung Association: Tobacco Policy Trend Alert, February 2007*, 1-8.
- Breslau, N., Johnson, E. O., Hiripi, E., & Kessler, R. (2001). Nicotine Dependence in the United States: Prevalence, Trends and Smoking Persistence. *Archives of General Psychiatry, 58*, 810-816.
- Chassin, L., Presson, C., Pitts, S., & Sherman, S. (2000). The natural history of cigarette smoking from adolescence to adulthood in a Midwestern common sample: Multiple trajectories and their psychosocial correlates. *Health Psychology, 19*(223-231).
- Delucchi, K. L., Matzger, H., & Weisner, C. (2004). Dependent and problem drinking over 5 years: a latent class growth analysis. *Drug and Alcohol Dependence, 74*(3), 235-244.
- Engeland, A., Haldorsen, T., Anderson, A., & Tretli, S. (1996). The Impact of Smoking Habits on Lung Cancer Risk: 28 Years' Observation of 26,000 Norwegian Men and Women. *Cancer Causes & Control, 7*(3), 366-376.
- Enstrom, J. E. (1999). Smoking Cessation and Mortality Trends Among Two United States Populations. *Journal of Clinical Epidemiology, 52*(9), 813-825.
- Freund, K., Belanger, A., D'Agostino, R., & Kannel, W. (1993). The health risks of smoking. The Framingham Study: 34 years of follow-up. *Annals of Epidemiology, 3*(4), 417-424.
- Freund, K. M. (1992). Predictors of smoking cessation: the Framingham Study. *American Journal of Epidemiology, 135*(9), 957-964.
- Giovannucci, E., Rimm, E.B., Stampfer M.J., Colditz, G.A., Ascherio, A., Kearney, J., Willett, W.C. (1994). A Prospective Study of Cigarette Smoking and Risk of Colorectal Adenoma and Colorectal Cancer in U.S. Men. *J Natl Cancer Inst, 86*, 183-191.
- Godtfredsen, N. S., Prescott, E., & Osler, M. (2005). Effect of Smoking Reduction on Lung Cancer Risk. *JAMA, 294*(12), 1505-1510.
- Gordon, T. K., William B., Dawber, T., McGee, D. (1975). Changes associated with quitting cigarette smoking: The Framingham Study. *The American Heart Journal, 90*(3), 322-328.

- Hill, J. A. (1936). United States Life Tables. *United States Department of Commerce: Bureau of the Census*.
- Hirdes, J. P., Brown, K. S., Vigoda, D. S., Forbes, W. F., & Crawford, L. (1987). Health Effects of Cigarette Smoking: Data from the Ontario Longitudinal Study on Aging. *Canadian Journal of Public Health, 78*(1), 13-17.
- Howard, G., Wagenknecht, L. E., Burke, G. L., Diez-Roux, A., Evans, G. W., McGovern, P., Nieto, J., Tell, G.S. (1998). Cigarette Smoking and Progression of Atherosclerosis. *JAMA, 279*(2), 119-124.
- Knoke, J. D., Shanks, T. G., Vaugh, J. W., Thun, M. J., & Burns, D. M. (2004). Lung Cancer Mortality Is Related to Age in Addition to Duration and Intensity of Cigarette Smoking: An Analysis of CPS-1 Data. *Cancer Epidemiology, Biomarkers & Prevention, 13*(6), 949-957.
- Lam, T. H., He, Y., Li, L. S., Li, L. S., He, F. S., & Liang, B. Q. (1997). Mortality Attributable to Cigarette Smoking in China. *Journal of the American Medical Association, 278*(18), 1505-1508.
- Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767-778.
- MMWR. (1993). Cigarette Smoking-Attributable Mortality and Years of Potential Life Lost - United States, 1990. *CDC Morbidity and Mortality Weekly Report, 42*(33), 645-648.
- MMWR. (2003). Cigarette Smoking-Attributable Morbidity --- United States, 2000. *CDC Morbidity and Mortality Weekly Report, 52*(35), 842-844.
- Methuén, B., & Methuén, L. K. (2000). Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol. Clin. Exp. Res, 24*(6), 882-891.
- Nagin, D. (1999). Analyzing developmental trajectories: a semi-parametric group-based approach. *Psychological Methods, 4*, 139-157.
- Nagin, D. (2005). *Group-Based Modeling of Development*. Cambridge, MA: Harvard University Press.
- NCHS. (2006). *Health, United States, 2006 With Chartbook on Trends in the Health of Americans*. Hyattsville, MD: National Center for Health Statistics.
- Nusselder, W. J., Looman, C. W. N., Mheen, P. J. M.-v. d., Mheen, H. v. d., & Mackenbach, J. P. (2000). Smoking and the compression of morbidity. *J Epidemiol Community Health 54*, 566-574.

- Ostbye, T., & Taylor, D. H. (2004). Smoking and Substance Abuse: The Effect of Smoking on Years of Health Life (YHL) Lost Among Middle-Aged and Older Americans. *HSR: Health Services Research*, 39(3), 531-552.
- Ostbye, T., Taylor, D. H., & Jung, S.-H. (2002). A Longitudinal Study of the Effects of Tobacco Smoking and Other Modifiable Risk Factors on Ill Health in Middle-Aged and Old Americans: Results from the Health and Retirement Study and Asset and Health Dynamics among the Oldest Old Survey. *Preventative Medicine*, 34, 334-345.
- Prescott, E., Hippe, M., Schnohr, P., Hein, H. O., & Vestbo, J. (1998). Smoking and risk of myocardial infarction in women and men: longitudinal population study. *BMJ*, 316, 1043-1047.
- Prescott, E., Osler, M., Hein, H. O., Borch-Johnsen, K., Schnohr, P., & Vestbo, J. (1998). Life expectancy in Danish women and men related to smoking habits: smoking may affect women more. *Journal of Epidemiol Community Health*, 52, 131-132.
- Rogot, E., & Murray, J. L. (1989). Smoking and Causes of Death Among U.S. Veterans: 16 Years of Observation. *Public Health Reports*, 95(3), 213-222.
- Simons, L. A., Simons, J., McCallum, J., & Friedlander, Y. (2005). Impact of smoking, diabetes and hypertension on survival time in the elderly: the Dubbo Study. *MJA*, 182(5), 219-222.
- Thun, M., Day-Lally, C., Calle, E., Flanders, W. D., & Heath Jr., C.W. (1995). Excess Mortality among Cigarette Smokers: Changes in a 20-Year Interval. *American Journal of Public Health*, 85(9), 1223-1230.
- Thun, M. J., & Heath Jr., C.W. (1997). Changes in Mortality from Smoking in Two American Cancer Society Prospective Studies since 1959. *Preventative Medicine*, 26, 422-426.
- USDHHS. (1990). The Health Benefits of Smoking Cessation. *U.S. Department of Health and Human Services: Public Health Service: Office on Smoking and Health*.
- USDHHS. (2004). The Health Consequences of Smoking: A Report of the Surgeon General. *U.S. 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*.
- Vaillant, G., Schnurr, P., Baron, J., & Gerber, P. (1991). A prospective study of the effects of cigarette smoking and alcohol abuse on mortality. *Journal of General Internal Medicine*, 6(4), 299-304.

- Vaillant, G. E. (1979). Natural History of Male Psychologic Health: Effects of Mental Health on Physical Health. *New England Journal of Medicine*, 301(23), 1249-1254.
- Vaillant, G. E. (1996). A Long-term Follow-up of Male Alcohol Abuse. *Arch. Gen. Psychiatry*, 53, 243-249.
- Vaillant, G. E. (1998). Natural History of Male Psychological Health, XIV: Relationship of Mood Disorder Vulnerability to Physical Health. *Am J Psychiatry*, 155(2), 184-191.
- Vaillant, G. E. (2003). A 60-year follow-up of alcoholic men. *Addiction*, 98, 1043-1051.
- Wannamethee, S. G., Shaper, A. G., & Perry, I. J. (2001). Smoking as a Modifiable Risk Factor for Type 2 Diabetes in Middle-Aged Men. *Diabetes Care*, 24(9), 1590-1595.
- Weir JM, D. J. J. (1970). Smoking and mortality: a prospective study. *Cancer*, 25(1), 105-112.
- Yuan, J.-M., Ross, R. K., Wang, X.-L., Gao, Y.-T., Henderson, B. E., & Yu, M. C. (1996). Morbidity and Mortality in Relation to Cigarette Smoking in Shanghai, China: A Prospective Male Cohort Study. *Journal of the American Medical Association*, 275(21), 1646-1650.