

CHILDHOOD CANCER SURVIVOR STUDY  
Analysis Concept Proposal

1. **TITLE:** Longitudinal smoking patterns in adult survivors of childhood cancer

2. **WORKING GROUP AND INVESTIGATORS**

Working group: This proposed study will be within the Cancer Control Working Group.  
Collaborating investigators will include (to date):

Todd M. Gibson	<a href="mailto:todd.gibson@stjude.org">todd.gibson@stjude.org</a>
Robert C. Klesges	<a href="mailto:rklesges@uthsc.edu">rklesges@uthsc.edu</a>
Greg Armstrong	<a href="mailto:greg.armstrong@stjude.org">greg.armstrong@stjude.org</a>
Melissa Hudson	<a href="mailto:melissa.hudson@stjude.org">melissa.hudson@stjude.org</a>
Deo Kumar Srivastava	<a href="mailto:kumar.srivastava@stjude.org">kumar.srivastava@stjude.org</a>
Wendy M. Leisenring	<a href="mailto:wleisenr@fhcrc.org">wleisenr@fhcrc.org</a>
Karen Emmons	<a href="mailto:kemmons@hsph.harvard.edu">kemmons@hsph.harvard.edu</a>
Ann Mertens	<a href="mailto:amerten@emory.edu">amerten@emory.edu</a>
Paul Nathan	<a href="mailto:paul.nathan@sickkids.ca">paul.nathan@sickkids.ca</a>
Les Robison	<a href="mailto:les.robison@stjude.org">les.robison@stjude.org</a>

3. **BACKGROUND AND RATIONALE**

Survival after diagnosis with a pediatric cancer has improved dramatically in recent decades, such that the five-year survival rate exceeds 80% [1]. Attendant with this improvement has been an increased emphasis on the long-term health of survivors, who maintain an elevated risk of mortality and morbidity many decades after their initial cancer treatment [2, 3]. Specifically, childhood cancer survivors have increased risk of second cancers, pulmonary complications, cardiac disease, and fertility impairments [4-8]. Smoking is an established risk factor for each of these outcomes in the general population, suggesting that childhood cancer survivors who smoke may be at particularly high risk. Therefore, it is essential to characterize smoking behaviors among childhood cancer survivors, and to identify opportunities for promotion of smoking prevention and cessation in this high risk group.

Numerous studies have examined smoking prevalence in childhood cancer survivors, and some have identified risk factors for smoking initiation or cessation [9-16]. Most recently, a study of 10,326 survivors in Britain reported lower prevalence of current and ever smoking in survivors (20% and 30%, respectively) compared to the comparable general population (28% and 49%, respectively) [17]. Among survivors, factors associated with being a current smoker included age at questionnaire completion (highest risk for ages 20-34), lack of radiation treatment, lower socioeconomic status, lower educational attainment, and lack of long-term hospital follow-up. Notably, a previous study by Emmons et al. examined predictors of smoking initiation and cessation among childhood cancer survivors using data from the baseline questionnaire of the CCSS [18]. In a retrospective cohort study of 9709 survivors, they found that 28% reported ever smoking ( $\geq 100$  cigarettes during lifetime) and 17% reported being current

smokers. Standardized to United States population rates, CCSS survivors were less-likely to be current smokers than the general population (standardized incidence ratio (SIR) = 0.72, 95% confidence interval (95% CI) = 0.69-0.75). Factors independently associated with smoking initiation included older age at cancer diagnosis, lower household income, less education, not having had pulmonary-related cancer treatment, and not having had brain radiation. Among ever smokers, smoking cessation prior to the baseline questionnaire was inversely associated with younger age at smoking initiation, less education, and having had brain radiation. Cessation was positively associated with very young age at cancer diagnosis (<3 years).

Although many studies have examined smoking in childhood cancer survivors, these studies have typically included data from only a single questionnaire, precluding them from detailing changes in smoking behaviors over time. The prevalence of smoking in the general population has decreased in recent decades in the United States [19], but it is unclear how this temporal trend has impacted smoking in childhood cancer survivors. At the individual level, identification of characteristics associated with long-term smoking patterns could inform prevention and control efforts.

The currently proposed study will extend the work of Emmons et al. by utilizing data from CCSS follow-up questionnaires to examine longitudinal patterns in smoking behaviors over time. However, such a longitudinal analysis must account for the Partnership for Health study (PFHS), an intervention trial conducted within the CCSS cohort from 1999-2001 [20]. The PFHS attempted to enroll all eligible current smokers from 22 of the 25 CCSS institutions, and ultimately enrolled 796 subjects (63% of potentially eligible). The randomized control trial found the peer-counseling group (intervention) had significantly elevated cessation rates compared to the self-help group (control) after 12 months [21], and that the higher cessation rates remained significant after 2-6 years of follow-up, although differences were not large (20.6% versus 17.6%) [22]. In the proposed study, cross-sectional estimates of smoking status will need to be interpreted in the context of this intervention carried out in a subset of the CCSS population. Although estimates can be made for the subset of the population not exposed to the intervention, results may not be generalizable to the full CCSS due to differences between survivors who did and did not participate in the PFHS.

Additionally, an ongoing analysis by Klesges et al (manuscript under review) examined the efficacy of tobacco quitlines in 519 regularly smoking survivors from the St. Jude Lifetime Cohort study, the After Completion of Therapy (ACT) Clinic, or the CCSS. Although some CCSS participants participated in this intervention study, recruitment started slowly and there is minimal overlap between the timeframe when participants completed the most recent CCSS questionnaire and initiation of recruitment into the trial. The intervention study found that a counselor-initiated telephone-based intervention resulted in significantly increased cessation rates at 8 weeks, but the effect was not maintained over 12 months. The proposed analysis will include investigation of how many participants were recruited for the quitline intervention study prior to returning the CCSS 2007 Follow-up questionnaire (F07), but this number is expected to be very low.

Given limited resources and competing priorities for health promotion efforts in childhood cancer survivors, it is important to target smoking prevention and cessation efforts at subgroups most likely to

benefit. Therefore, we propose an examination of longitudinal smoking patterns to identify factors associated with the long-term smoking experience of survivors.

#### **4. SPECIFIC AIMS/OBJECTIVES/RESEARCH HYPOTHESES**

**Objective:** To assess prevalence of smoking status in CCSS over time (cross-sectional), and to identify and characterize longitudinal patterns in smoking behaviors among adult childhood cancer survivors who completed both baseline and follow-up questionnaires.

**Aim 1.** Describe the prevalence of smoking status (current, former, never) in the CCSS population over calendar time based on each of the three questionnaires (baseline, F03, F07). This will include assessment of potential selection bias due to differential loss to follow-up over time (as described in the Analysis Plan). Compare this with a) the prevalence of smoking status in CCSS siblings at the same time points and b) prevalence in an age, sex, and race/ethnicity-matched general population sample.

Hypothesis: The prevalence of current smoking will decrease over time in both survivors and siblings, and will be below that expected in the general population at all time intervals.

**Aim 1a.** In current smokers at baseline with available follow-up data, identify the proportion who report being former smokers at each follow-up questionnaire among a) those who did not participate in Partnership for Health, b) those who enrolled in the Partnership for Health intervention group, and c) those who enrolled in the Partnership for Health control group (i.e. intention to treat populations). The goal of this aim is not to determine the effectiveness of the intervention, which has been reported previously, but rather to provide context for the prevalence estimates presented in Aim 1 (i.e. a subset of the survivors enrolled in an intervention trial, and Aim 1a provides an indication of how this may or may not have impacted the overall CCSS prevalence estimates at each time point).

Hypothesis: Current smokers in the CCSS who did not participate in the Partnership for Health study will have smoking cessation rates lower than those in either the peer-counseling intervention group or the self-help control group, but differences will be small by the 2007 questionnaire.

**Aim 2.** Identify the prevalence and predictors of specific longitudinal patterns in smoking behaviors over time (e.g. consistent current smoker, consistent never smoker, long-term quitter). Examine associations between demographic, cancer-related, or health status variables (at baseline and changes over time for time-dependent variables) and membership in a certain longitudinal pattern subgroup (Brinkman et al, 2013 [23]; see Analysis Framework for details on subgroups).

Hypothesis: Education, socioeconomic status, age at initiation, and prevalence of chronic conditions will be associated with specific longitudinal smoking patterns.

**Aim 3:** Characterize the cumulative exposure of CCSS participants to tobacco smoking, based on the most recently available follow-up data. Among those with valid self-reported ages of initiation and/or cessation and smoking intensity over all three questionnaires, calculate cumulative exposure as pack-years of smoking. Determine pack-years of exposure for current smokers in F07, ever smokers in the cohort, and by specified risk groups (e.g. cancer diagnosis, pulmonary or cardio-toxic therapy, age at initiation, prevalence of chronic conditions).

Hypothesis: Cumulative exposure will be lower for those with pulmonary toxic therapy or chronic conditions, as well as higher age at initiation.

## 5. Analysis Framework

**Population:** For Aim 1, the population for each time point will include all survivors and siblings who completed a CCSS questionnaire (excluding proxy respondents) at that time point. Analyses of period prevalence in the adult CCSS population at each questionnaire will examine all survivors or siblings age  $\geq 18$  at the time of that questionnaire (i.e. excluding parent-reported exposures). Therefore some participants excluded from the baseline analysis will be included in the analysis of F03 and F07 data after reaching age 18. Analyses of longitudinal patterns in smoking over time will be restricted to participants with smoking status self-reported from at least two different questionnaires. Smoking status in the general population will be estimated using data from the National Health Interview Survey (NHIS), a large-scale household interview survey of a statistically representative sample of the U.S. civilian non-institutionalized population (<http://www.cdc.gov/nchs/nhis.htm>).

### Primary Exposures:

1. Smoking status: never/former/current/missing  
Baseline(BL): N.1 ( $\geq 100$  cigarettes in entire life, yes/no) and N.1.d (Do you smoke now, y/n)  
F03: L.1 ( $\geq 100$  cigarettes in entire life, y/n) and L.2 (Do you smoke now, y/n)  
F07: N.7 ( $\geq 100$  cigarettes in previous 2 years, y/n) and N.9. (Do you smoke now, y/n)  
For status at F07, a “yes” for N.7 will be a former or current smoker, depending on N.9.  
If “no” for N.7, responses at F03 and BL need to be considered to distinguish never/former
2. Longitudinal smoking patterns: 1. persistent smoker, 2. consistent nonsmoker, 3. long-term quitter, 4. relapser, 5. late initiator

See Appendix 1 for exposure matrix

3. Pack-years: (Total years smoked) x (Average # cigarettes per day)  
To determine the best estimate at F07, we need to also incorporate some data from F03 (see Analysis Plan for details).  
F03: L1. ( $\geq 100$  cigarettes in entire life), L3. (average cigarettes per day) and L4. (total years smoked).  
F07: N7. ( $\geq 100$  cigarettes in previous 2 years), N10. (average cigarettes per day) and N11. (total years smoked).

**Secondary Exposures:**

1. Quit proportion: a) proportion of population classified as former smokers, among all those with smoking status at that time interval  
b) proportion of smokers from previous time interval now classified as former (i.e. % quit during the interval)
2. Smoking intensity: Among the subset of persistent current smokers, examine the number of cigarettes per day at each time interval (BL: N.1b; F03: L.3; F07: N.10)

**Covariates:**

1. Sex
2. Race/Ethnicity
3. Age at diagnosis
4. Age at initiation
5. Radiation treatment (cranial, chest, other)
6. Chemotherapy (agents with suspected cardio or pulmonary toxicities)
7. Grade 1-2 chronic conditions
8. Grade 3-4 chronic conditions
9. Attained age
10. Childhood cancer diagnosis
11. Education
12. Income (baseline and change over time)
13. Marital status (baseline and change over time)
14. Other tobacco use (y/n; baseline and change over time)

**Analysis Plan:****Aim 1:**

**Part 1:** Describe the baseline characteristics (Table 1: age at diagnosis, current age, race, gender, smoking status, initial cancer diagnosis, education/SES, treatment categories) of survivors according to their participation status in a) CCSS follow-up questionnaires from 2003 (F03) and 2007 (F07), and b) the Partnership for Health smoking cessation intervention study (which included only current smokers at baseline who agreed to participate). It will be important to show these comparisons so that readers can understand the potential impact of a) attrition bias due to loss to follow-up or b) selection bias due to participation in the intervention trial (i.e. if we examine patterns in those not exposed to the intervention, they may not be representative of smokers in CCSS overall). We will use chi-square statistics to test for heterogeneity in baseline characteristics between respondents of each follow-up questionnaire and the baseline group. Differences will be described and incorporated into data interpretation as appropriate. Similarly, we will compare characteristics among current smokers at baseline who enrolled in the Partnership for Health Study and those who were not enrolled, including examination of potential differential participation in future CCSS follow-up questionnaires (by study enrollment or intervention assignment).

**Part 2:** Determine the proportion of survivors with determined smoking status at baseline (i.e. age  $\geq 18$ , answers to N1 and N1.d) who self-reported as never smokers, former smokers, or current smokers. Determine the proportions of baseline smoking status among siblings with available data (i.e. age  $\geq 18$ , answers to N1 and N1.d) at the baseline questionnaire, and describe stratum-specific prevalence if siblings differ significantly from survivors on age, sex, or race/ethnicity. Estimate smoking status proportions in the general population using data from an age, sex, and race/ethnicity-matched sample from the 1993 National Health Interview Survey (matched to the survivor sample, to the extent this is possible with level of detail and sample size of the NHIS datasets; we will explore the possibility of adjusting for geographic region as well). Graph these proportions as shown in Figure 1. Repeat these calculations among all survivors and siblings with available data at F03 and F07. For the general population comparison to the F03 data, use NHIS 2003 data, and for F07 data use NHIS 2007. The core smoking status questions based on having ever smoked  $\geq 100$  cigarettes and level of current smoking are consistent across these NHIS questionnaires.

**Aim 1a:** Examine only the group of participants who self-reported as current smokers at baseline. Among those with available smoking status at F03 or F07, determine the proportion who self-reported as being a former smoker at each time interval, stratified by whether they were enrolled in the Partnership for Health Study intervention group, the Partnership for Health study control group, or not enrolled in the Partnership for Health Study. Non-enrollees are likely to have lower follow-up rates compared to enrollees, and among enrollees follow-up may be different between those assigned to intervention or control (as will be captured in Table 2 above). We are interested in the proportion of CCSS respondents at each time interval who reported being quitters at that interval. If there are substantial differences in quit rates by Partnership for Health participation status, then differential follow-up between these groups may contribute to observed differences over time. Note: the goal is not to determine the unbiased effect of the intervention study, but to see whether the existence of the intervention study may have altered overall CCSS prevalence estimates. Logistic regression (or log-linear) models will be used to compare quit proportions across time points, with adjustment for age, sex, and race/ethnicity. We will explore the possibility of adjusting for Partnership for Health status if necessary.

**Aim 2.** Assign CCSS participants with smoking status available for at least two questionnaires to one of six longitudinal patterns described in Appendix 1. The prevalence proportion of each pattern will be determined based on the percentage of all participants eligible to be assigned a pattern (i.e. two or more data points) that fall into each individual pattern. The five *a priori* defined patterns of interest will be a) consistent nonsmokers, b) persistent current smokers, c) long-term quitters, d) relapsers, and e) late initiators. Participants not falling into any of these patterns will be included in a sixth “Other” pattern. Definitions of these patterns are as follows, as reflected in the exposure matrix in Appendix 1 (all with  $\geq 2$  data points, which requires assumptions in the case of missing data) or Appendix 2 (if we restrict to only those who answered all three questionnaires):

Consistent nonsmoker: reported as “never smoker” on all questionnaires

Persistent current smoker: reported as “current smoker” on all questionnaires

Long-term quitter: reported as “former smoker” on at least two questionnaires, with no reports of “current smoker” at F03 or F07

Relapser: reported as “former smoker” and then subsequently as “current smoker” (note: due to wording in F07, participants could be “former” at F07 and still have reported initiating smoking in the last two years (i.e. started and stopped again), in which case they would be categorized as a “relapser”).

Late initiator: reported as “former” or “current smoker” after at least one report of “never smoker”

\* Participants with inconsistent longitudinal patterns (e.g. current – never – never) will be excluded.

Polytomous logistic regression models will be used to examine associations between baseline patient characteristics or cancer-related variables (shown in Table 1) and membership in a particular longitudinal pattern. We will also examine the associations with changes in status over time for household income, marital status, and chronic conditions (i.e. changes occurring during the interval where longitudinal patterns were defined). Each variable will be assessed separately in a univariate model in an initial exploratory analysis, but the primary analysis will identify significant predictors of pattern membership in a multivariable model. Due to potential challenges in interpreting individual regression coefficients in polytomous logistic models, we may choose to present predicted marginal probabilities rather than relative risk ratios, or may elect to fit standard logistic regression models where each smoking pattern is evaluated versus a single reference pattern (e.g. consistent nonsmokers).

**Aim 3:** Estimate the cumulative exposure to cigarette smoking among participants who answered all three questionnaires. We will use the most recent available self-reported smoking data, ignoring discrepancies with previous reports (i.e. when pack-years differ using data from F03 and F07, we will use the F07 data). However, due to the wording of question N7 on F07, determining the best estimate of cumulative exposure in some participants will require incorporation of data from F03. For participants reporting not having smoked  $\geq 100$  cigarettes in the last two years on F07, we will estimate cumulative exposure based on the response to the F03 questionnaire. A limitation of this approach is the chance that we would miss exposure in participants who started smoking after F03 and then quit more than 2 years before F07. However, it seems reasonable to conclude that this would involve a small number of participants. We will calculate pack-years for a) persistent current smokers (i.e. current at all three questionnaires) and b) ever smokers. We will also compare pack-years among ever smokers by subgroups defined by a) cancer diagnosis, b) pulmonary or cardio-toxic treatment received, c) prevalence of chronic conditions, and d) age at initiation. Differential length of follow-up between groups will be minimized by including only those participants who answered all three questionnaires.

## 6. References

1. Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, *et al.* SEER Cancer Statistics Review, 1975-2010. In. Bethesda, MD: National Cancer Institute; [http://seer.cancer.gov/csr/1975\\_2010/](http://seer.cancer.gov/csr/1975_2010/), based on November 2012 SEER data submission, posted to the SEER web site, April 2013
2. Armstrong GT, Liu Q, Yasui Y, Neglia JP, Leisenring W, Robison LL, *et al.* Late mortality among 5-year survivors of childhood cancer: a summary from the Childhood Cancer Survivor Study. *J Clin Oncol* 2009,**27**:2328-2338.
3. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, *et al.* Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med* 2006,**355**:1572-1582.
4. Friedman DL, Whitton J, Leisenring W, Mertens AC, Hammond S, Stovall M, *et al.* Subsequent neoplasms in 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. *J Natl Cancer Inst* 2010,**102**:1083-1095.
5. Green DM, Kawashima T, Stovall M, Leisenring W, Sklar CA, Mertens AC, *et al.* Fertility of female survivors of childhood cancer: a report from the childhood cancer survivor study. *J Clin Oncol* 2009,**27**:2677-2685.
6. Green DM, Kawashima T, Stovall M, Leisenring W, Sklar CA, Mertens AC, *et al.* Fertility of male survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *J Clin Oncol* 2010,**28**:332-339.
7. Mertens AC, Yasui Y, Liu Y, Stovall M, Hutchinson R, Ginsberg J, *et al.* Pulmonary complications in survivors of childhood and adolescent cancer. A report from the Childhood Cancer Survivor Study. *Cancer* 2002,**95**:2431-2441.
8. Mulrooney DA, Yeazel MW, Kawashima T, Mertens AC, Mitby P, Stovall M, *et al.* Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study cohort. *BMJ* 2009,**339**:b4606.
9. Bellizzi KM, Rowland JH, Jeffery DD, McNeel T. Health behaviors of cancer survivors: examining opportunities for cancer control intervention. *J Clin Oncol* 2005,**23**:8884-8893.
10. Demark-Wahnefried W, Werner C, Clipp EC, Guill AB, Bonner M, Jones LW, *et al.* Survivors of childhood cancer and their guardians. *Cancer* 2005,**103**:2171-2180.
11. Haupt R, Byrne J, Connelly RR, Mostow EN, Austin DF, Holmes GR, *et al.* Smoking habits in survivors of childhood and adolescent cancer. *Med Pediatr Oncol* 1992,**20**:301-306.
12. Larcombe I, Mott M, Hunt L. Lifestyle behaviours of young adult survivors of childhood cancer. *Br J Cancer* 2002,**87**:1204-1209.
13. Mulhern RK, Tyc VL, Phipps S, Crom D, Barclay D, Greenwald C, *et al.* Health-related behaviors of survivors of childhood cancer. *Med Pediatr Oncol* 1995,**25**:159-165.
14. Tao ML, Guo MD, Weiss R, Byrne J, Mills JL, Robison LL, *et al.* Smoking in adult survivors of childhood acute lymphoblastic leukemia. *J Natl Cancer Inst* 1998,**90**:219-225.
15. Tyc VL, Hadley W, Crockett G. Predictors of intentions to use tobacco among adolescent survivors of cancer. *J Pediatr Psychol* 2001,**26**:117-121.
16. Verrill JR, Schafer J, Vannatta K, Noll RB. Aggression, antisocial behavior, and substance abuse in survivors of pediatric cancer: possible protective effects of cancer and its treatment. *J Pediatr Psychol* 2000,**25**:493-502.
17. Frobisher C, Winter DL, Lancashire ER, Reulen RC, Taylor AJ, Eiser C, *et al.* Extent of smoking and age at initiation of smoking among adult survivors of childhood cancer in Britain. *J Natl Cancer Inst* 2008,**100**:1068-1081.



18. Emmons K, Li FP, Whitton J, Mertens AC, Hutchinson R, Diller L, *et al.* Predictors of smoking initiation and cessation among childhood cancer survivors: a report from the childhood cancer survivor study. *J Clin Oncol* 2002,**20**:1608-1616.
19. Pierce JP, Messer K, White MM, Cowling DW, Thomas DP. Prevalence of heavy smoking in California and the United States, 1965-2007. *JAMA* 2011,**305**:1106-1112.
20. Emmons KM, Butterfield RM, Puleo E, Park ER, Mertens A, Gritz ER, *et al.* Smoking among participants in the childhood cancer survivors cohort: the Partnership for Health Study. *J Clin Oncol* 2003,**21**:189-196.
21. Emmons KM, Puleo E, Park E, Gritz ER, Butterfield RM, Weeks JC, *et al.* Peer-delivered smoking counseling for childhood cancer survivors increases rate of cessation: the partnership for health study. *J Clin Oncol* 2005,**23**:6516-6523.
22. Emmons KM, Puleo E, Mertens A, Gritz ER, Diller L, Li FP. Long-term smoking cessation outcomes among childhood cancer survivors in the Partnership for Health Study. *J Clin Oncol* 2009,**27**:52-60.
23. Brinkman TM, Zhu L, Zeltzer LK, Recklitis CJ, Kimberg C, Zhang N, *et al.* Longitudinal patterns of psychological distress in adult survivors of childhood cancer. *Br J Cancer* 2013,**109**:1373-1381.

**Appendix 1. Proposed exposure matrix for longitudinal smoking patterns, with inclusion of participants with one missing data point.**

Baseline smoking status	F03 smoking status	F07 smoking status	Pattern
Never	Never	Never	Consistent nonsmokers
Never	Never	Missing	Consistent nonsmokers
Never	Missing	Never	Consistent nonsmokers
Missing	Never	Never	Consistent nonsmokers
Current	Current	Current	Persistent smokers
Current	Current	Missing	Persistent smokers
Current	Missing	Current	Persistent smokers
Missing	Current	Current	Persistent smokers
Current	Former	Former (<100 last 2yrs)	Long term quitter
Former	Former	Former (<100 last 2 yrs)	Long term quitter
Former	Missing	Former (<100 last 2 yrs)	Long term quitter
Former	Former	Missing	Long term quitter
Missing	Former	Former (<100 last 2 yrs)	Long term quitter
Former	Current	any	Relapser
Former	Former	≥100 last 2 years*	Relapser
Missing	Former	≥100 last 2 years*	Relapser
Former	Missing	≥100 last 2 years*	Relapser
Current	Former	≥100 last 2 years*	Relapser
Never	Never	Current	Late initiator
Never	Never	Former	Late initiator
Never	Current or Former	Any (exclude Never)	Late initiator
Missing	Never	Current	Late initiator
Missing	Never	Former	Late initiator
Never	Missing	Current or Former	Late initiator

\* In F07, a participant categorized as “former” may have initiated smoking in the past two years but subsequently reported quitting again, so we can’t use the strict definition of “current”

**Appendix 2. Proposed exposure matrix for longitudinal smoking patterns, including only those participants with smoking status at all three time intervals.**

Baseline smoking status	F03 smoking status	F07 smoking status	Pattern
Never	Never	Never	Consistent nonsmokers
Current	Current	Current	Persistent smokers
Former	Former	Former	Long term quitter
Former	Current	any	Relapser
Former	Former	≥100 last 2 years*	Relapser
Current	Former	≥100 last 2 years*	Relapser
Never	Never	Current	Late initiator
Never	Never	Former	Late initiator
Never	Current or Former	Any (exclude Never)	Late initiator

Table 1. Baseline characteristics of CCSS participants with self-reported smoking status data

Note: We will also evaluate characteristics of siblings and those who filled out the questionnaire but have missing smoking data, and will report these if warranted.

	Baseline		Follow-up 1		<i>p</i> *	Follow-up 2		<i>p</i> *
	N	%	N	%		N	%	
Total								
Age at baseline (years)								
<18								
18-20								
21-24								
25-29								
30-34								
35-39								
40-49								
50-59								
60+								
Race/Ethnicity								
White (non-Hispanic)								
Black (non-Hispanic)								
Hispanic								
Other								
Not Specified								
Household Income (\$/year)								
< 20,000								
20,000-59,999								
60,000+								
Not specified								
Marital Status								
Single								
Married/living as								
Widowed								
Divorced/separated								
Not specified								
Educational status								
1-12 years (not HS grad)								
High school grad								
College grad, postgrad								
Not specified								
Cancer diagnosis								
Leukemia								
CNS								
Hodgkin's								
NHL								
Kidney (Wilms)								
Neuroblastoma								
Soft tissue sarcoma								
Bone cancer								
Age at cancer diagnosis (years)								
0-5								
5-10								
10-15								
16-20								
Cancer treatment								
Cranial radiation								
Chest radiation								
Cardio/pulmonary-toxic chemo								

\* Test for heterogeneity compared with baseline

Table 2. Baseline characteristics of participants who reported current smoking at baseline

Note: Baseline characteristics between the intervention and control arm of the trial are similar due to randomization, but we will evaluate and report the proportions in each arm that completed the two CCSS follow-up questionnaires to evaluate differential follow-up.

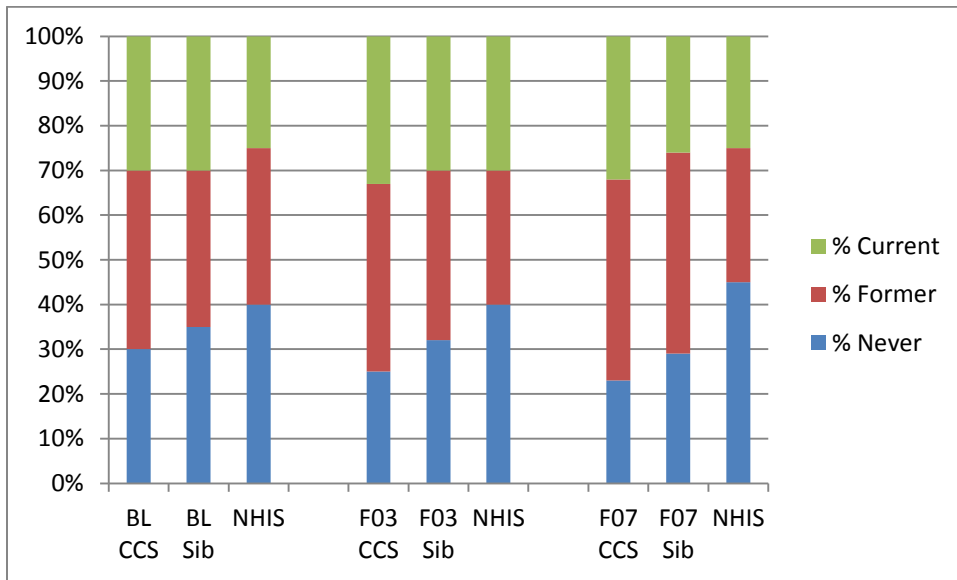
	Partnership for Health Smoking Intervention Trial				<i>P</i> *
	Enrolled		Not Enrolled		
	N	%	N	%	
Total					
Smoking status available for FU 1					
Smoking status available for FU2					
Age at baseline (years)					
<18					
18-20					
21-24					
25-29					
30-34					
35-39					
40-49					
50-59					
60+					
Race					
White (non-Hispanic)					
Black (non-Hispanic)					
Hispanic					
Other					
Not Specified					
Household Income (\$/year)					
< 20,000					
20,000-59,999					
60,000+					
Not specified					
Marital Status					
Single					
Married/living as					
Widowed					
Divorced/separated					
Not specified					
Educational status					
1-12 years (not HS grad)					
High school grad					
College grad, postgrad					
Not specified					
Cancer diagnosis					
Leukemia					
CNS					
Hodgkin's					
NHL					
Kidney (Wilms)					
Neuroblastoma					
Soft tissue sarcoma					
Bone cancer					
Age at cancer diagnosis (years)					
0-5					
5-10					
10-15					
16-20					
Cancer treatment					
Cranial radiation					
Chest radiation					
Cardio/pulmonary-toxic chemo					

\* Test for heterogeneity

Note: Characteristics other than data availability in follow-up questionnaires have been reported by Emmons et al.

Figure 1. Smoking status prevalence among questionnaire respondents, age 18 or older at time of questionnaire

**Sample data for illustration:**



Note 1: NHIS = prevalence in the age/race/sex-matched general population in the same time period

Note 2: The proportion of survivor and sibling respondents with missing smoking status will also be reported

Figure 2. Percent of current smokers at baseline who reported being former smokers at follow-up (this figure shows the potential impact of the intervention study on the CCS prevalence trends reported in Figure 1)

**Sample data for illustration:**

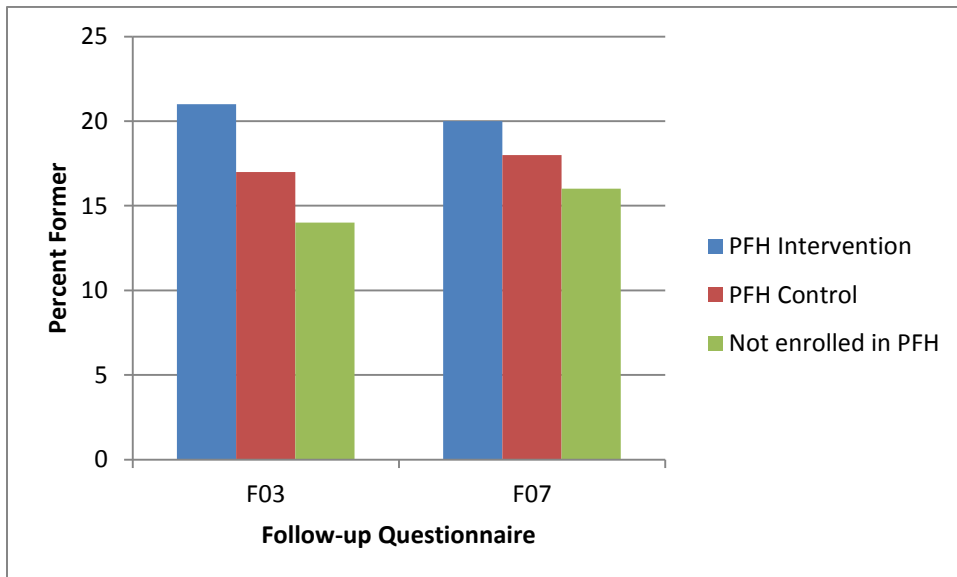


Table 3. Univariate polytomous logistic regression models predicting longitudinal class membership

	Consistent never smoker N = xx (%)	Persistent current smoker N = xx (%)	Long-term quitter N = xx (%)	Relapser N = xx (%)	Late initiator N = xx (%)	Other pattern N = xx (%)
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Female sex	1.00 (reference)					
Young age at diagnosis	1.00 (reference)					
Older attained age	1.00 (reference)					
Non-white race/ethnicity	1.00 (reference)					
Low household income at BL	1.00 (reference)					
Decrease in income	1.00 (reference)					
Married status at baseline	1.00 (reference)					
Change in marital status	1.00 (reference)					
Education status	1.00 (reference)					
Cranial radiation	1.00 (reference)					
Chest radiation	1.00 (reference)					
Cardio/pulmonary-toxic chemotherapy	1.00 (reference)					
Chronic conditions at BL						
Mild/moderate	1.00 (reference)					
Severe/disabling	1.00 (reference)					
Increased chronic conditions						
Mild/moderate	1.00 (reference)					
Severe/disabling	1.00 (reference)					

Note: Age and income categories to be determined

Table 4. Multivariable polytomous logistic regression model predicting longitudinal class membership

	Consistent never smoker N = xx (%)	Persistent current smoker N = xx (%)	Long-term quitter N = xx (%)	Relapser N = xx (%)	Late initiator N = xx (%)	Other pattern N = xx (%)
	RRR (95% CI)	RRR (95% CI)	RRR (95% CI)	RRR (95% CI)	RRR (95% CI)	RRR (95% CI)
Variable A	1.00 (reference)					
Variable B	1.00 (reference)					
Variable C	1.00 (reference)					
Variable D	1.00 (reference)					
Variable E	1.00 (reference)					

\* Depending on ease of interpretation of results, we may either present marginal shifts in probabilities due to change in covariate values, or fit separate logistic regression models to illustrate associations

Table 5 or Figure 3. To be determined based on feasibility and initial analyses of Aim 3 (cumulative exposure in pack-years)