Study 1 Observation Study
Mechanisms of Ethnic/Racial Differences in Lung Cancer Due to Cigarette Smoking
part of the Clinical and Biomarkers Core

CLINICAL PROTOCOL

CPRC # 2017NLS019

Principal Investigators:

Dorothy Hatsukami, Ph.D.
UMN Department of Psychiatry

Loic Le Marchand, MD, PhD
University of Hawaii Cancer Center

Irina Stepanov, Ph.D.*
Division of Environmental Health Sciences

Charles Rosser, MD
UH Cancer Center

*non-clinical - will not consent participants

Biostatistician:

Daniel Stram, Ph.D.
University of Southern California

Version 4
Date: July 2, 2020

Revision History

Revision #	Version Date	Summary Of Changes	Consent Changes
1	1-20-2017	original to CPRC-IRB - rejected	
2	4-17-2017	Revised for Study 1 only	
3	8-10-2017	Removed reference to investigational product	No
4	7-2-2020	COVID-19 modifications to reduce contact and enforce social distancing • Discontinue carbon monoxide • Some of the questionnaires done at home • Addition of COVID-19 screening questions • Temperature taken before visit • Use of PPE during visits Compensation changed to $75	Yes

Principal Investigator Contact Information:
Dorothy Hatsukami, Ph.D.
717 Delaware St SE #260
hatsu001@umn.edu
612-626-2121

Table of Contents

Table of Contents

Abbreviations

AE: **Adverse Events**: is an undesired harmful effect resulting from a medication or other intervention / study procedure.

BP: **Blood Pressure:** is the pressure exerted by circulating blood upon the walls of blood vessels and is one of the principal vital signs.

CYP2A6: This gene encodes a member of the cytochrome P450 superfamily of enzymes. The cytochrome P450 proteins are monooxygenases which catalyze many reactions involved in drug metabolism. The enzyme is known to metabolize nicotine and nitrosamines. Individuals with certain allelic variants are said to have a poor metabolizer phenotype, meaning they do not efficiently metabolize nicotine.

FTND: **Fagerstrom Test for Nicotine Dependence:** is a 6-item standard instrument for assessing the intensity of physical addiction to nicotine and includes an evaluation of cigarette consumption, the compulsion to use, and dependence.

HR: **Heart Rate:** a measure of the number of heart beats per minute (bpm)

MEC **Multiethnic Cohort**

NIAAA: **National Institute on Alcohol Abuse and Alcoholism:** part of the U.S. National Institutes of Health that supports and conducts biomedical and behavioral research on the causes, consequences, treatment, and prevention of alcoholism and alcohol-related problems.

NMR: **Nicotine Metabolite Ratio:** is a urinary measure of the ratio of nicotine metabolites, which indicates speed of nicotine metabolism.

NNAL: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol: Human exposure to NNK can be measured by analyzing the sum of NNAL and its glucuronides (total NNAL) in urine.

NNK 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone: One of the tobacco specific nitrosamines, a potent lung carcinogen in tobacco products, which are formed from nicotine during the curing and processing of tobacco.

NNN: **N'-nitrosonornicotine:** One of the tobacco specific nitrosamines formed from nicotine during the curing and processing of tobacco and a known carcinogen.

PAH: **Polycyclic Aromatic Hydrocarbons:** PAH, a product of combustion, are mong the most important carcinogens in tobacco smoke.

SAE: **Serious Adverse Events:** generally, any event which causes death, permanent damage, birth defects, or requires hospitalization is considered an SAE.

TLFB: **Timeline Follow-Back:** is a method that can be used as a clinical and research tool to obtain a variety of quantitative estimates of marijuana, cigarette, and other drug use by asking clients to retrospectively estimate their usage prior to the interview date.

TNE: **Total nicotine equivalents:** a urinary measure of nicotine and its metabolite concentrations.

Study Synopsis

The Clinical and Biomarkers Core of the Program Project entitled "Mechanisms of Ethnic/Racial Differences in Lung Cancer Due to Cigarette Smoking" will recruit smokers in Hawaii to collect biological samples and will carry out analyses of tobacco constituent biomarkers for four projects. The biological samples collected will be used to evaluate and compare biomarkers of tobacco exposure across Japanese Americans, Whites, and Native Hawaiians on multiple urinary or blood components to assess tobacco related exposures.

The samples will either be procured from the existing biorepository of the Multiethnic Cohort MEC Study for life-long non-smokers or obtained from smokers who are recruited from the MEC or general population by this Core. **This protocol includes the scope of work for the Clinical and Biomarkers Core related to recruitment of human subjects (Observation Study) for the 3 projects that will use and/or store the biorepository samples.** Biorepository samples will be available to the PPG research team or other qualified investigators examining issues related to cancer and/or tobacco exposure biomarkers.

Study Design: The Clinical and Biomarkers Core will include two components: the Clinical Recruitment and Data Collection component and the Biomarker Analysis component. The purpose of the Clinical Recruitment. Data Collection component is to provide human biological samples and measurements from people of various ethnic and racial backgrounds for Projects 2, 3 and 4 (Project 1 is conducted at the University of Southern California and University of Hawaii on samples from the MEC Biorepository, Southern Community Cohort Study (SCCS), Singapore Cohort Study, Shanghai Chinese Health Study and Atherosclerosis Risk in Communities (ARIC). This Core provides a centralized approach to subject recruitment, biological sample collection, and analysis of validated biomarkers applying well-established protocols and using standardized methodologies. The subject population is recruited from the Multiethnic Cohort Study (MEC) and the general population. The samples collected under this protocol will be used to evaluate and compare biomarkers of tobacco exposure across Japanese Americans, Whites, and Native Hawaiians and to establish a biorepository to develop or assess future biomarkers. In addition, for Project 2, we will conduct a short clinical study with a subsample of Japanese Americans using deuterium-labeled tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) cigarettes for NNK metabolic profiling

This protocol is for Study 1 – Observation Study.

Primary Aim: The Clinical and Biomarkers Core will facilitate the research by collecting biological samples and biomarker data from 300 smokers that will be used by individual projects to achieve their Specific Aims.

Patient Population: Cigarette smokers who are part of the targeted ethnic/racial groups including Japanese Americans, Non-Hispanic Whites and Native Hawai'ians. Subjects will be recruited from male and female participants of the Multiethnic Cohort (MEC) Study or the general population (recruited through local advertisements).

Accrual Goal: Years 1 -3: 100 subjects in the Observation Study.

Enrollment Period: September 1, 2017 through December 31, 2020

1 Study Objectives

1.1 Primary Objectives for Projects Requiring Recruited Subject Samples*

- Project 2: Identify and recruit a subset of Japanese smokers (N=20 selection based on slow and fast metabolizers based on CYP2A6 activity) to determine the effect of CYP2A6 genotype on NNK metabolic activation and will be conducted under a separate Study 2: Clinical Study protocol.
- Project 3: Obtain biomarker samples and observational data on 300 Native Hawaiian, White, and Japanese American smokers (100 in each group) to investigate ethnic differences in butadiene-DNA adducts and determine the relationship of DNA adducts to lung cancer and the influence of carcinogen metabolizing genes on DNA adduct formation, repair, and toxicity/mutagenicity.
- Project 4: Obtain oral mucosa cell samples on 300 Native Hawaiian, White, and Japanese American smokers to quantify DNA adducts and investigate urinary biomarkers of acrolein, crotonaldehyde, inflammation, and oxidative damage among smokers and non-smokers.

(*Project 1 will use existing samples from the existing biorepository and are not part of this Core).

2 Background and Significance

The major goal of this project is to provide insights into the underlying mechanisms of ethnic/ racial differences in lung cancer due to cigarette smoking. In a previous study, we carried out extensive carcinogen metabolite phenotyping studies, genome wide association studies (GWAS), and mechanistic investigations to test the hypothesis that these differences in lung cancer susceptibility are due to ethnic and racial variations in exposure and response to cigarette smoke carcinogens.

Multiethnic Cohort Study (MEC)

The MEC is a prospective cohort study of 215,251 men and women between the ages of 45 to 75 at baseline, primarily belonging to five racial/ethnic groups: African American, Japanese American, Latinos, Non-Hispanic, and whites. This cohort was established to investigate the association of lifestyle and genetics with chronic diseases in a multiethnic population. [61] Between 1993 and 1996, potential participants were identified in Hawai'i and California (primarily Los Angeles County) through drivers' license files, voter registration lists, and Health Care Financing Administration files. Each participant completed a mailed, self-administered questionnaire regarding demographic, dietary, lifestyle, and other exposure factors at baseline and follow-up questionnaires updating smoking and other exposures every five years.

Tobacco constituent biomarkers to be assessed:

Nicotine is the major known addictive constituent of cigarette smoke and its uptake determines smoking behaviors in individual smokers.[1] Since each dose of nicotine inhaled with cigarette smoke is accompanied by numerous toxic and carcinogenic compounds,[2] nicotine uptake is critical in determining carcinogen dose in smokers. Total nicotine equivalents (TNE) is the sum of urinary nicotine, nicotine-N-oxide, cotinine, and 3'-hydroxycotinine and their glucuronides, which together account for 73 – 96% of the nicotine dose.[1] This biomarker is an outstanding monitor of nicotine uptake and level of smoking.[1] Nicotine metabolite ratios can also provide other important information, such as differences in glucuronidation and in cytochrome P450 2A6 activity. The ratio of total 3'-hydroxycotinine to cotinine will be used for this purpose in this program.[3,4]

4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) is a potent lung carcinogen in every species tested, inducing lung tumors independent of the route of administration.[5,6] Based on animal data, mechanistic studies, and epidemiological evidence, the International Agency for Research on Cancer (IARC) classified NNK and the related carcinogen NNN as human carcinogens.[7] Human exposure to NNK can be measured by analyzing the sum of NNAL and its glucuronides (total NNAL) in urine.[8-10] Total NNAL correlates with cigarettes per day, nicotine equivalents in urine, and cotinine,[6,11] and is an effective uptake marker for NNK.[12] Consistent with the known carcinogenicity of NNK, three nested case-control studies in smokers have demonstrated that NNAL is related to lung cancer.[13-15] Thus, NNAL is both an exposure and a risk biomarker.

Polycyclic aromatic hydrocarbons (PAH) are among the most important carcinogens in tobacco smoke.[16] Phenanthrene tetraol (PheT) is a metabolite of the non-carcinogenic PAH phenanthrene, which is the simplest polycyclic aromatic hydrocarbon (PAH) with a “bay region”, a structural feature associated with carcinogenicity.[17-19] PheT forms through the diol epoxide metabolic pathway of phenanthrene, mimicking the carcinogenic pathway of the prototypic carcinogenic PAH benzo[*a*]pyrene (Figure 2).[20-26] The end product of this pathway for phenanthrene is PheT in urine, which can serve as a biomarker of PAH exposure and metabolic activation.[23] PheT levels correlate with those of 1-hydroxypyrene, an established biomarker of PAH exposure,[22] and are 2-3 times higher in smokers than in non-smokers.[24] Urinary PheT is related to lung cancer in a nested case-control study in male smokers.[15] Another phenanthrene metabolite excreted in urine is 3-hydroxy phenanthrene (3-PheOH); this biomarker represents PAH exposure and detoxification.[27] The ratio PheT:3-PheOH can therefore be a phenotypic indicator of an individual’s ability to metabolically activate.

Volatile compounds such as acrolein and crotonaldehyde may also contribute to lung cancer induction in smokers.[28] Acrolein is a highly toxic agent in cigarette smoke, causing cilia-toxicity among other effects in the lung.[29] One study has shown that acrolein causes a pattern of DNA damage in the p53 tumor supressor gene similar to that observed in lung tumors of smokers.[30] We have demonstrated the presence of substantial quantities of acrolein-DNA adducts in human lung DNA.[31] Crotonaldehyde, a structural homologue of acrolein, also forms DNA adducts in the human lung, and these are known to be mutagenic, although crotonaldehyde has not been shown to be carcinogenic to the lung.[32] Acrolein and crotonaldehyde are detoxified by reactions with glutathione, and the resulting conjugates are excreted in the urine as 3-hydroxypropylmercapturic acid (3-HPMA) and 3-hydroxy-1-methylpropylmercapturic acid (HMPMA), respectively. These mercapturic acids can be readily measured by established and analytically validated methods, and are related to cigarette smoking.[33-36] Previous research shows levels of these mercapturic acids were unusually elevated in Native Hawaiians, an aspect that will be further investigated in Project 4.

Cadmium (Cd) is a human lung carcinogen, with cigarette smoking being one of the major sources of exposure.[37] Urinary Cd increases with age and smoking, as major determinants of its levels.[10] A statistically significant association between urinary Cd levels and lung cancer mortality in men from the NHANES III cohort has been reported.[38] While there is some overlap in urinary levels of Cd between smokers and non-smokers, the levels of this biomarker have been demonstrated to increase with smoking.[10,39,40] Exposure to Cd has been shown to cause hypermethylation of tumor suppressor genes[41] and was inversely associated with LINE1 methylation, which is a common epigenetic event in carcinogenesis.[42]

Oxidative stress, inflammation, and related biomarkers

Chronic inflammation and oxidative stress are important interlinked contributing factors in the pathogenesis of cigarette smoke-associated diseases, including lung cancer.[43-48] Oxidative

damage is likely involved in tumor promotion and other deleterious effects of smoking (reviewed in [49,50]). Cigarette smoke contains high levels of prooxidants, such as reactive oxygen and nitrogen species, as well as catechols and related compounds which can induce oxidative damage and lead to inflammatory state.[51,52] Inflammation involves infiltration of lymphocytes, macrophages, and neutrophils into tissues under stress, and induces lipid peroxidation, these processes resulting in the generation of a spectrum of reactive oxygen and nitrogen species capable of causing extensive damage to DNA and proteins, and resulting in toxic and mutagenic events.[53,54] In smokers, inflammation and oxidative stress will be assessed by measuring urinary metabolites and DNA adducts.

Inflammation leads to the induction of the immediate-early response gene prostaglandin G/H synthase-2 (COX-2).[55] The product of COX-2, prostaglandin E_2, is metabolized to PGEM, a urinary metabolite and promising biomarker of inflammation which has been associated with cancer development in a variety of studies.[55] Levels of PGEM are reported to be significantly higher in current smokers than in former smokers than in never-smokers, which is thought to reflect inflammation in the lung due to induction of COX-2.[56,57] Oxidative stress-induced peroxidation of arachidonic acid results in the production of F_2-iosprostanes, among which urinary 8-*iso*-PGF2α is an accepted biomarker of oxidative damage.[58,59] Many studies consistently report that levels of 8-*iso*-$PGF_{2\alpha}$ are elevated in smokers (reviewed in [39,60-62]). In DNA, guanine is the major target for the direct oxidation by the inflammation-induced radicals. The oxidation produces several DNA adducts including 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxo-dG), which can lead to chromosomal aberrations and the induction of mutations, and is widely used as biomarker for oxidative stress.[63]

Clinical and Biomarkers Core's Role

The Clinical and Biomarkers Core will provide services, biomarker samples and tools for all projects in this program project grant to test new hypotheses put forth and will provide new insights into the ethnic/racial differences in lung cancer risk due to smoking.

3 Rationale

Smoking-related lung cancer is the leading cause of cancer death in the United States and worldwide. For a given lifetime exposure to cigarette smoking, the risk of lung cancer is greatest in Native Hawaiians and lowest in Japanese Americans compared to Whites. The results of this program project will provide unique and critical data relevant to the mechanism(s) underlying the high risk of Native Hawaiian smokers for lung cancer; thus, potentially generating new insights for lung cancer prevention strategies. Using a biomarker approach, previous research found a direct link of tobacco carcinogens NNK and PAH to lung cancer, and NNN to esophageal cancer in humans. These novel findings are strongly consistent with the results from experimental animal models and implicate the carcinogenic role of these specific tobacco constituents in the development of lung and esophageal cancers. This project will build on these results and further provide rationale to regulate and control these constituents in tobacco products to reduce the harmful effects of tobacco smoking.

The purpose of Study 1: Observation Study under the Clinical and Biomarkers Core of the "Mechanisms of Ethnic/Racial Differences in Lung Cancer Due to Cigarette Smoking" program project is to provide human biological samples and measurements from people of various ethnic and racial backgrounds for Projects 2, 3 and 4. These samples will be used to evaluate and compare biomarkers of tobacco exposure across Japanese Americans, Whites, and Native Hawaiians and to add to the Multiethnic Cohort MEC biorepository to develop or assess future biomarkers.

Clinical and Biomarker Core Project Aims

Project 1 in this program will investigate the relationship of biomarkers of tobacco constituent exposure and early-biological effects with the risk of lung cancer. MEC and other pre-existing samples will be used for this analysis.

Project 2 will focus on characterization of NNK bioactivation across different ethnic groups and the effect of P450 2A6 activity on this process. This requires recruitment of smokers from three ethnic/racial groups (Native Hawaiians, Whites, and Japanese Americans) for the analysis of biomarkers of NNK α-hydroxylation (bioactivation) pathway, as well as recruitment of a subgroup of Japanese American smokers for participation in a clinical study in which deuterium-labeled NNK will be used to specifically trace NNK-derived metabolites. The Clinical and Biomarkers Core will carry out subject recruitment and the clinical study.

Project 3 will investigate the role of 1,3-butadiene metabolic activation and deactivation in lung cancer risk among various ethnic/racial groups. This project will require urine samples from smokers and nonsmokers from the three ethnic/racial groups recruited by the Clinical and Biomarker Core for the analysis of 1,3-butadiene DNA adducts. Data on nicotine intake (urinary TNE) in these subjects as well as in 400 lung cancer cases and 400 controls from Project 1 will be also required for this project.

Project 4 will investigate whether DNA adduct levels in oral cells, themselves or together with urinary biomarker levels, correlate with lung cancer risk across various ethnic/racial groups. This project will require biological samples from smokers and nonsmokers from these ethnic/racial groups, as well as data on urinary tobacco constituent biomarkers and PGEM and 8-*iso*-$PGF_{2\alpha}$, in these subjects.

This protocol will recruit smokers, and will generate the biomarker data required. The Coordinating site at the University of Minnesota will not recruit subjects, but provide oversight for the study team recruiting subjects at the University of Hawai'i.

4 Study Design Overview

This cross-sectional, observational study will primarily recruit smokers from the MEC study. The biological samples from these smokers are being collected since they do not currently exist in the biorepository (e.g. buccal cells). Potential smokers from the three racial groups who meet specific inclusion criteria will be invited to participate in the study. Should recruitment from the MEC sample be insufficient, smokers who meet the criteria will be recruited from the general population. This study will involve one to two home visits where tobacco use and medical history and biological samples will be collected including blood, buccal cells and urine.

5 Participant Selection

Approximately 300 smokers will be recruited from the Multiethnic Cohort (MEC) study or the general population in Hawai'i. 100 participants from each of the three designated ethnic/racial groups with equal distribution by gender. All participants will be recruited from Hawai'i to reduce variation due to geographical location and maximize our ability to recruit the targeted ethnic/racial groups. Based on the data collected from the MEC, participants who are at least 21 years old and who currently smoke greater than 5 cigarettes per day (current smokers) will be selected and contacted in the MEC standard method by sending a letter inviting them to participate in this study. The letter states that a researcher will be in contact with them to provide further information and obtain verbal consent to submit to the telephone screening questionnaire, if interested. In the MEC database, the approximate number of Oahu participants that might be available for each of the three categories of participants is listed in Table 2. From past experience with Oahu MEC participants, including smokers[64] we expect a participation rate of ~50%. Several hundreds of additional MEC current smokers are also available from the neighboring islands.

Table 2. Number of potential Oahu MEC participants available for each category.

Race/ethnicity	Smokers*
Japanese Americans	810
Non-Hispanic Whites	397
Native Hawaiians	385

* As determined from the most recent follow-up questionnaire

Recruitment beyond the MEC sample will be expanded to the general population of smokers in Hawai'i. Recruitment will conducted via local print, radio, TV, or internet advertisements and posted flyers.

5.1 Inclusion Criteria

a) One of the three targeted ethnic groups:
   - Japanese American - two parents of Japanese descent
   - Non-Hispanic Whites - two parents of non-Hispanic white descent
   - Native Hawaiians will include individuals with at least one parent of Hawaiian descent;

b) Smoke 5 cigarettes per day over the past three months;

c) ≥21 years of age;

d) Consumes 14 or fewer drinks of alcohol per week;
e) Generally stable and good health (determined by review of medical history);
f) Able to provide written voluntary consent before performance of any study related procedure.

5.2 Exclusion Criterion

a) Current use of other nicotine containing products for > 4 times per month (and no use of any nicotine-containing products except cigarettes for 2 weeks prior to their study visits);
b) Acute or uncontrolled medical or psychiatric conditions;
c) Currently taking any medications that affect relevant metabolic enzymes or anti-inflammatory medications such as ibuprofen (this will be reviewed by study investigators on a case-by-case basis);
d) Active infection (e.g., influenza, cold, respiratory infection, sinus infection) at the time of the visit;
e) Pregnant or nursing or planning on becoming pregnant during the study;
f) Unable to read and understand English.

6 Participant Registration

Registration will occur after the subject consent form is signed and eligibility is confirmed, but before any study samples or data are collected. To be eligible for registration to this study, the participant must meet each criteria listed on the eligibility checklist based on the documented eligibility assessment.

Registration with the Masonic Cancer Center Clinical Trials Office

Upon completion of obtaining written consent, screening evaluation, and eligibility confirmation, the study coordinator or designee will enroll the subject and enter registration information into the study registration system. This data will be provided to the Masonic Cancer Center.

7 Study Procedures for Recruitment:

7.1 Observation Study

Participants recruited from the MEC pool who meet eligibility will be sent a letter informing them of the study. After receipt of the letter, staff will contact the subject and obtain verbal consent for telephone screening to further assess eligibility. If the subject is interested and eligible they will be scheduled for an in-home visit or a visit in the clinic, whichever is preferred. Subjects recruited from the general population will contact the clinic in response to advertisements to be screened over the telephone for eligibility.

7.1.1 Observation Study-Visit 1 Procedures: (see Table 3 Procedures by Visit)

Visits 1 and 2 may be combined if biological samples can be collected at Visit 1. During the first visit, potential participants will be more fully informed about the study and written consent will be obtained.

Prior to any visit, COVID-19 symptoms or potential exposure to COVID-19 will be assessed. Prior to entering the clinic, participants will have their temperature taken and asked to wear a mask. Staff will be using appropriate PPE during the visits.

Once informed consent is obtained, more extensive and updated information will be obtained. The following measures will be reviewed by the researcher to assist in confirming eligibility:

Questionnaires completed by participant at home or clinic or by interview:

1. Tobacco Use History, Tobacco Dependence and Exposure Interview.
2. Brief medical history.
3. Medication use (current and last 30 days).
4. NIAA Alcohol Use Questionnaire (12 month and 1 month)
5. Timeline Follow-back assessing tobacco, nicotine, other combustibles or alcohol use for last 14 days.

Physiological Measures:

1. Blood pressure.
2. Heart rate.
3. Body weight (in kilograms).
4. Height.
5. Calculate body mass index and possible body fat composition.

Scheduling Visit 2

At the end of Visit 1, the second visit will be scheduled approximately 1 week after Visit 1 (+7 to 14 days). The interviewer will provide the subject with:

1. A urine container for first morning void sample.
2. A specimen collection commode (female participants only).
3. A cooler and two packs of blue-ice.
4. Instructions to freeze the blue-ice in advance of the collection, collect the first morning urine void upon rising the morning of the second visit, and store the sample in the cooler on the blue-ice.
5. **Instructions to avoid grilling foods two days before the biomarker sample collection visit (to avoid exposure to PAH in cooked meats).**
6. Smokers will be reminded not to use smoked products (e.g., cigars, marijuana, etc.) except cigarettes before the next visit.
7. Avoid any other tobacco/nicotine product, including electronic cigarettes prior to the visit.

7.1.2 Observation Study-Visit 2 Procedures (Visit +7 to 14 days)

Visit 1 and 2 are typically combined.

At the second visit, the following exposure questionnaires completed by participant at home or in the clinic or by interview will be administered:

1. Environmental Tobacco Smoke Questionnaire.
2. Environmental and Occupational Exposure History (including volcanic fumes).
3. Grilling, smoking or barbecuing meat or fish in the past 48 hours.
4. Timeline Follow-Back assessing tobacco, nicotine, other combustibles or alcohol and drug use since Visit 1.
5. Medication and nutritional supplements (including name, dose per pill, number of pills and indication for use) will be obtained by visual inspection of the products.

Collection of biological samples.

1. First morning void urine samples will be collected for analysis of biomarkers.
2. Blood sample (non-fasting): Three tubes of blood (2- 10mL lavender EDTA tubes and 1 10 mL red top tube) will be drawn by trained staff. If the participant is

seen in their home, the phlebotomist visit may occur on a separate day due to scheduling concerns.

3. Oral cell samples collected.

7.1.3 Compensation:

Participants will be compensated $75 per visit.

8 Subjective and Exposure Forms for Observation Study

8.1 Questionnaires:

a. Demographics questionnaire will assess age, gender, race, ethnicity, income, education, current occupation and usual occupation.
b. Tobacco Use Exposure and History Interview assesses variables such as number of cigarettes smoked per day, age of initiation of smoking, duration of quit attempts, duration of smoking, pack years and other tobacco product use.
c. The Fagerstrom Test for Nicotine Dependence is the most widely used and psychometrically tested scale for dependence.[65]
d. Medical History and Current Health Status list current diagnoses, symptoms and past health problems.
e. Concomitant Medications includes any medications taken within 30 days prior to the first visit and throughout study participation in Project 2.
f. NIAAA Alcohol Use Questionnaire examines rate of alcohol use for the past month.[67-68]
g. Environmental Tobacco Smoke Questionnaire consists of 6 questions related to smoke exposure at home, work and socially.[69]
h. Environmental and Occupational Health History assesses environmental and occupational toxicant exposures, which might affect our biomarker measures.[70]
i. 48 hour Food Recall will be administered by trained study staff to assess ingestion of any foods that have might affect the biomarkers that are assessed (e.g., grilled foods or smoked meats, cruciferous vegetables).
j. Adverse Effects Checklist lists several nicotine exposure and toxicity and respiratory symptoms related to smoking.
k. Health Changes Questionnaire.
l. Timeline Follow-Back.
m. Daily Diary card provides daily number of cigarettes, other tobacco/nicotine products and number of alcoholic drinks per day.

9 Laboratory Procedures for processing of biological samples:

All collection containers will be labeled with study number, subject number, collection date and type of biological sample (urine, blood, buccal cells).

9.1 Blood Collection and Processing

Blood collected into purple top tubes will be processed within 12 hours of phlebotomy. The tubes will be centrifuged at 1,200 × g for 10 minutes in a refrigerated centrifuge to separate plasma, buffy coat, and red blood cells. The separated components will be transferred into cryogenic vials, labeled (with information described above) and stored in a -80°C freezer. The urine samples will be aliquoted into cryotubes and frozen at -70°C. Tubes with mouthwash and buccal cell samples will be centrifuged at 5000 x g for 15 minutes to pellet the cells and the supernatant will be discarded. The cell pellets will be washed by adding 1 mL of 1X TE buffer (pH 7.4), swirling the tube gently, repeating the centrifugation, and discarding the buffer. The washed buccal cell pellet will be re-suspended in 200 µL cold 1X TE buffer, transferred into a pre-labeled cryovial, and stored at -80°C until isolation of DNA

for adduct analyses in Project 4.

9.2 Oral Cell Collection and Processing

Participants will be asked to brush their teeth at the beginning of the visit by using a provided commercial individually packed pre-pasted toothbrush (ReadyBrushTM) and asked to not eat, drink, smoke or chew gum for 30 minutes prior to obtaining oral cells. Two types of oral cell collection methods will be used on each subject: an oral rinse with mouthwash and buccal cell scrapings. Sample collection with mouthwash will be carried out first; participants will be provided with 10 mL of mouthwash or saline with approximately in a sterile 50-mL polypropylene tube and asked to swish the mouthwash throughout the mouth for 45 seconds. The swishing will be performed in the following manner: 15 seconds on the left side, 15 seconds on the right side, and 15 seconds throughout the mouth, including the mouth floor and the area between the lips and gums. This technique will ensure the most complete collection of exfoliated cells from different parts of the oral cavity, and minimize possible variations in the efficiency of sample collection by different subjects. The mouthwash sample will be expectorated back into the 50-mL polypropylene tube. Buccal cells will be collected by scraping the oral mucosa inside the mouth with a cytobrush. Two cytobrushes will be used on each subject to collect two separate samples from the left and the right inner cheek. After scraping the mucosal surface, each cytobrush will be immediately placed into a separate pre-labeled sterile polypropylene tube containing 5 mL of Scope mouthwash, swirled for several seconds, and tapped against the walls of the tube to ensure the cells are transferred from the brush to the Scope. Samples will be stored in a -20 freezer.

9.3 Urine Collection and Processing

Subjects are instructed to collect their first morning void on the day of the visit and keep it in the provided cooler and on freezer packs. Staff will collect the sample and verify with the subject when the sample was collected. Urine will be aliquoted into 4.5 mL cryo vials labeled with study specific scanner labels upon returning to the laboratory and stored in a -20C freezer. Vials will be stored by assay type and shipped upon request.

9.4 Biomarker Management Website

Throughout the study, biological samples (blood, buccal cells, urine) will be collected, labeled and registered in the University of Hawai'i's MEC Tracking Website. Samples will be sent in split batches to the University of Southern California (USC) Biomarker Core for processing and aliquots sent to Masonic Cancer Center, University of Minnesota, to be stored in the Biorepository with de-identified information. Some samples will remain with the PI at the University of Hawaii. The UHCC MEC, USC Biomarker Core and Masonic Cancer Center Tobacco Biorepositories have established policies and procedures consistent with NCI guidelines. Samples will be stored until fully used, no longer usable, or destroyed upon subject request. A discussion of the storage, future use, and sharing of samples is contained in the informed consent for this study. A Data and Biospecimen Sharing Plan is also provided in section 17.

10 Clinical Data Collection

Sources of materials:

The primary source of clinical data will be the protocol specific case report forms and a web database. Subjective questionnaires will be completed by the subject directly into the internet based database that is password protected and meets HIPAA standards. All data from the subjects will be on a secure web-site or sent to the University of Minnesota, and all information will be de-identified. All biosamples sent to the University of Minnesota will also be de-identified and location and disposition will be kept on the Biomarker Management Website housed on the Masonic Cancer Center server.

For the Observational Study, subjects will be asked to provide information on their tobacco use history, medical history, medication use and other measures that might affect biomarkers such as alcohol or other drug use, food, secondhand smoke exposure and occupational history. Biological samples including urine and oral cells will be collected.

All records will be confidential. All shared data will be de-identified. University of Hawaii Cancer Center personnel will ensure all the study records are accurate and that all consents have been signed, thereby having access to names of individuals. University of Minnesota staff will monitor progress through a web-based tracking database and regular research phone calls. University of Minnesota staff may also visit Hawaii to review protocol adherence and forms. However, no record of this identifying information will leave the premises. Identifying information will be kept locked in a secure area.

Human subject consent forms will indicate corresponding subject ID and will not be stored with the data. Paper and pencil subjective forms, and other Case Report Forms (CRF) will be identified with subject number and initials, and visit number and date will be entered onto appropriate forms. Data entered into the study database and will be de-identified, with only subject number as a link. Biomarker samples will be labeled with a scanner barcode indicating subject number, visit number, specimen type (e.g., urine, serum, saliva) and intended assay or biobank samples.

11 Risks of Study Participation

Potential risks: The potential risks for subjects are minimal. For both studies, most physiological and subjective measures will be noninvasive and should present no psychological or medical risk to the subject. Some questionnaires may be of a sensitive nature assessing subjects' alcohol use. Subjects will be told that they may refuse to answer, however, refusal may effect continued participation in the study.

In the observational study, subjects will be using their own products so no additional product risk is imposed. The additional risk in this study is loss of privacy due to a breach in confidentiality.

Protection against Risks

Potential subjects will be told the nature of the research over the phone and during screening. Subjects will be required to demonstrate an understanding of the study purpose and procedures prior to signing the consent form. This consent form must be signed before the research is started. They will be told they may discontinue participation at any time and will not be discriminated against if they choose to do so and that discontinuing this study will not affect their further participation as a MEC participant.

Protection against breach of confidentiality:
Subjects will be told their participation in the project will be strictly confidential, that any identifying information will only be available to the site investigators, coordinating center study manager, an auditor for the Food and Drug Administration, study sponsor or University of Hawai'i's institutional oversight. No identifying information concerning the data and results will be made known. Subjects will have written assurance that while de-identified individual subject data may be available to other researchers for research purposes, only a summary of the results will ever be published or otherwise publicly released. They will also be informed that all raw data will be coded with numbers and any identifiers will be kept in locked file cabinets and only appropriate study personnel will have access.

All data will be de-identified and transferred through a secure, password-protected website that is only available to the investigators. All identifying information will be in a locked in a

secure place and accessible only to relevant study staff.

Risk with phlebotomy:
A trained laboratory technician will obtain blood samples to minimize the risk associated with a blood draw. Blood samples will be obtained by a trained phlebotomist with single-use, sterile blood collection supplies.

Throughout their participation in the study, adverse symptoms will be recorded at each clinic visit and monitored by the PIs and study medical personnel. Subjects who experience any significant adverse events will be sent or accompanied to the emergency clinic or referred to their physician, depending on the nature and severity of the adverse event.

12 Effect on Patient Care

This study has no effect on patient care.

13 Duration of Study Participation

Participation will consist of one to two visits in the Observation Study over a one day to three week period (if two visits are conducted, they will be approximately one week apart as the preferred time frame, however we will accommodate the participant schedule as needed).

14 Adverse Event Reporting

This study carries minimal risk; however it will comply with the University of Minnesota and University of Hawai'i IRB reporting requirements. Events requiring prompt reporting include any adverse event that requires a change to the protocol or consent form, any unauthorized disclosure of confidential information, any unresolved subject complaint or any protocol deviation resulting in harm or the unanticipated death of an enrolled subject. All study related adverse events of a non-serious nature will be reported to each institution's IRB at the time of IRB renewal application submission.

Reportable adverse events or unanticipated problems will be reported to the IRB by phone within 24 hours of when the investigator becomes aware of the event and written report within 10 days of receipt of information regarding the event at UH and submitted within 5 business days at UMN. The Data and Safety Monitoring Board will review all serious or unexpected adverse events and provide recommendations.

Any event requiring prompt reporting to the IRB will also be reported to the Masonic Cancer Center's SAE Coordinator, University of Hawai'i Cancer Center and the funding agency.

15 Study Data Collection and Monitoring

15.1 Data Management

This study will use a tracking database to follow the progress of each subject's participation. Patient demographics and study specific data will be entered into a database maintained by the UH Cancer Center.

15.2 Case Report Forms

Participant data will be collected using protocol specific electronic case report forms and pen and paper questionnaires. The case report forms will be developed and purchased by the coordinating site and distributed to the University of Hawaii site. The case report form books will be created to ensure capture of all data points. Each visit will have a checklist of all the measures that need to be taken and the order by which these measures are administered.

15.3 Data and Safety Monitoring Plan (DSMP)

The study coordinator and the supervising physician (Charles Rosser, MD) at the site of data collection will be responsible for the daily oversight of subject safety. Entrance criteria will be reviewed following screening. Subjects will be under supervision while in the study and seen on an ongoing basis by our research staff who will assess adverse events and vital signs and make appropriate referrals to the study physician.

Dr. Hatsukami and Joni Jensen will conduct a study initiation meeting and provide internal monitoring of subject safety. If any safety issues are identified, the investigator and staff at the University of Hawaii will be contacted immediately.

A Data and Safety Monitoring Board will not be implemented for this observational study.

15.4 Record Retention

The investigator will retain study records including source data, copies of case report form, consent forms, HIPAA authorizations, and all study correspondence in a secured facility for at least 6 years after the study file is closed with the IRB and FDA.

16 Statistical Considerations

Listed below are the projects using this protocol's collected samples. Statistical considerations related to Project 1 are found in the protocol related to that study.

Project 2: NNK α –hydroxyl glucuronides, metabolic profiling and activation
Aim 1: is to determine the relationship between CYP2A6 activity and NNK □hydroxylation among two groups of Japanese smokers one with very low measured CYP2A6 activity and the other with normal levels (about 10 per group). Our comparison will be between the means (or if the distribution is highly skewed the medians or geometric means) of the two groups (null versus average CYP2A6). The power of the test depends on the size of the group difference compared to the within-group variability, i.e. to the standard deviation of NNK- hydroxylation..Assuming that a normal distribution holds we will have 80 percent of the distribution of NNK α –hydroxylation (33 percent of variation explained). In these analyses we will adjust for potential confounders including age, sex, race, total nicotine equivalents, BMI, etc.

Aim 2: will test for ethnic differences in α –hydroxyNNK glucuronides among Japanese-Americans, Native Hawaiians, and white smokers recruited in Core B. In addition the glucuronides will be related to total urinary NNL and NNK DNA adducts analyzed in Project 4 (Aims 1 and 2). We will use all 300 smokers recruited by the Clinical Studies Core (100 in each of three ethnic groups). For these smokers α –hydroxyNNK glucuronides will be measured in this project and related to NNAL, TNE, and NNK DNA adducts measured by Project 4. For each of the glucuronides we will examine the data for outliers and failed QC values, evaluate batch differences, heteroscedasticity and the appropriateness of log or other transformation to achieve normality of the outcome measures. Linear regression methods will be used to characterize the association between (possibly transformed) glucuronides and the variables (NNAL, TNE, NNK adducts) of interest. We will adjust for age, sex, BMI and ethnicity, and examine independent effects of NNAL or NNK adducts with and without adjusting for TNE. With 300 smokers having data available we have 80 percent power to detect a variance explained of approximately 3 percent.

Aim 3: This is a laboratory aim (which will develop a NNK metabolic profiling method) and does not specifically test hypotheses needing statistical analysis.

Project 3: Ethnic/Racial Differences in metabolism and DNA Adduct Formation by 1,3-Butadiene

Aim 1: This Aim concentrates upon a subset of the participants recruited in the Clinical Core (the 100 smokers from each of three ethnic groups) with the objective being to characterize differences in metabolites of 1,3butadiene as well as DNA adduct formation by race/ethnicity. Levels of metabolites and of DNA adducts of 1,3-Butadiene will be compared between Japanese-Americans, Native Hawaiians, and non-Hispanic Whites. Adjustments for age, sex, TNE, CPD, and BMI will be performed using the "LS means" option in Proc GLM in SAS. Prior to model-fitting we will examine QC metrics and estimate laboratory variation in the outcomes based on duplicate samples from the same individual and time period. We will identify batch effects and individual outliers. If required, polynomial or log transformations will be used to normalize the distribution of metabolite and DNA adduct levels prior to analysis. With 100 smokers from each of the three groups we will have 80 percent power to detect a difference between any two ethnic groups in these outcomes which amounts to approximately 6 percent of the total variance of the outcome. Here we correct (using the Bonferroni method) for making 3 such comparisons by requiring $p<.017$ in order for any single comparison to be declared significant.

Aim 2: This aim investigates the association between urinary butadiene DNA adducts and lung cancer risk among White and African-American participants (smokers) in project 1. One hundred cases and one hundred controls for each ethnic group will be assessed. Controls will be frequency matched to the cases by age and sex. The primary issue is whether DNA adducts convey independent information beyond that which is contained in smoking dose, either self-reported (CPD) or measured (TNE). We will utilize unconditional logistic regression to evaluate the association between risk and DNA adducts after including CPD and TNE into the models. With 200 cases and 200 controls total (pooling both ethnicities) in this analysis we will be able to detect a 2.25 fold lung cancer risk difference between the highest and lowest quintiles of the distribution of DNA adducts using a test for log-linear trend. Effects specific to a single ethnic group must be larger (3-fold between top and bottom quintiles) than this to be detectable.

Aim 3 of this project is to relate inter-individual differences in metabolic inactivation, DNA adduct formation/repair and toxicity/mutagenicity of butadiene-derived epoxides measured in 40 Hapmap cell lines first to GSTT1 and EPHX1 activity level (low versus high for each) with the cell lines chosen as follows:

Cell lines number	GSTT1 activity	EPHX1 activity
10	High	High
10	High	Low
10	Low	High
10	Low	Low

In addition we will investigate the relationship between these phenotypes and genetic polymorphisms in specific candidate genes for carcinogen metabolism and repair. There are several important issues to be dealt with in this project by the Statistical Core. These include (1) implementing procedures for summarizing mutagenicity and toxicity for each of the 40 cell lines, and (2) selection of an appropriate power or log transformation that will normalize the distribution of the estimated values.

For the genetic analyses we chose the HapMap cell lines because of the extensive genotyping and other phenotyping (expression array studies etc.) already performed for these samples. We will only use founders (no offspring will be selected). Here we attempt to elate the cytotoxic, genotoxic and DNA repair endpoints to the genotypes available for these same cell lines (www.hapmap.org). For this problem, we simply fit

R1: Yi = a + b1Gi + ei

where Gi represents the main effects of a particular genetic marker (genotyped for individual i) Yi can represent any one of the cytotoxic, genotoxic, or DNA adduct/repair phenotypes.

Power:

For the comparisons of metabolites, adducts, etc. by GSTT1 and EPHX1 activity level and assuming no interactions between the two activities we will be able to detect with 20 individuals per group approximately a 1.2 standard deviation difference between the means of each phenotype. In our power calculations for the genetic analyses of Aim 3 we distinguish between analyses of candidate markers within a small number of candidate genes compared to the candidate markers and analyses of all 4 million markers genotyped in the HapMap. Our search for genetic markers related to the phenotypes of interest will be done in two stages. The first stage will focus only on SNPs in or near specific candidate genes, i.e. those genes previously identified as being highly likely to be related to each phenotype. It is expected that approximately 100 candidate SNPs will be identified for each phenotype triggered by each activated carcinogen. For the second stage all (approximately 4 million) SNPs will be considered in the performance of a genome-wide association scan for each phenotype. The statistical methods for these two stages are identical (based on model R1). The only difference will be the type I error rate at which candidate SNPs will be considered to be globally significant versus the remainder of the SNPs. For the analysis of candidate SNPs we will use a p-value for significance of 0.0005 for the 100 candidate SNPs and 2.5 x 10-8 for genome wide significance. The second p-value is based on assuming, as described by Pe'er et al [1], that the testing burden for testing all common SNPs in the HapMap in a single experiment is equivalent to performing approximately 1 million tests. For the candidate genes we will have approximately 80 percent power to detect HapMap SNPs which explain approximately 38 percent of the variance of a given phenotype after correcting for all associations tested. For the remainder of the SNPs only extremely strong associations (63 percent of variance explained) will be detectable at the genome wide level of significance.

Project 4: Oral cell DNA adducts and urinary biomarkers to investigate ethnic/racial differences in lung cancer susceptibility.

Aim 1: Evaluation of ethnic differences in DNA adducts in oral cells quantified in 100 smokers from Core B in each of three ethnic groups. This aim will quantify DNA adducts of acrolein, crotonaldehyde, NNN, NNK, and formaldehyde and identify ethnic differences in levels of these adducts. After QC, outlier identification, analysis of duplicate samples for laboratory variation and batch effects we will fit linear regression models to the (possibly log or power-transformed) adduct levels to allow for differences by age, sex, BMI, TNE and CPD while examining ethnic group differences. The SAS procedure GLM will be used to provide adjusted means (LS means) and to test for the significance of the observed ethnic differences in those means.

As in Project 3, Aim 1, for a given outcome (DNA adduct) we will have 80 percent power to detect a difference between any two ethnic groups which amounts to approximately 6 percent of the total variance of a given outcome. We can also examine the joint effects of smoking and ethnicity in the total of 300 Core B participants and will have 80 percent power to detect ethnicity x smoking interactions that explain at least an additional 1 percent of variation for any outcome.

Aim 2: There are two sub aims to Aim 2. The first is to compare urinary metabolites in both smokers and nonsmokers in the three ethnic groups. The main hypothesis is that acrolein and crotonaldehyde metabolites, as well as metabolites from oxidative damage and

inflammation are higher in Native Hawaiian non-smokers than in the other 3 ethnic groups, and that this is the reason that acrolein and crotonaldehyde metabolites were higher in Native Hawaiian smokers than Japanese American smokers in our published study. We will also compare the levels of the urinary metabolites in the smokers from the three ethnic groups. We will also compare smokers to non-smokers from each ethnic group (for metabolites of acrolein, and crotonaldehyde and biomarkers of oxidative damage and inflammation). Sub aim 2 is to relate DNA adducts measured for the 300 participants in Core B to urinary metabolites of TNE, total NNAL, mercapturic acids of acrolein, and crotonaldehyde measured in the same individuals. After QC, examinations of batch effects and laboratory variation, outlier detection, transformation to normality, we will estimate the ethnic differences and provide linear regression estimates of the relationships between the DNA adducts and metabolites. For sub aim 1 we will compare the urinary metabolite levels in each group of 100 with the other groups. We will use a p-value for any pairwise comparison of 0.003 as protection against type I error. For sub aim 2 since we anticipate that many different analyses will be performed (each DNA adduct against each relevant metabolite) we will use a Bonferroni-corrected p-value of 0.05/50 = 0.001 to declare significance in these analyses. Detectable effects will correspond to approximately 3 percent of variance of the DNA adducts explained.

Aim 3: Identification and quantification of known and previously unknown DNA adducts. This aim has two sub-aims. The first is to identify known and unknown DNA adducts related to smoking status among 50 current smokers and never smokers of single ethnic group (Whites). The second sub aim is to determine ethnic differences in DNA adduct levels found to be elevated in smokers in the first sub-aim. DNA oral cell adducts will be measured in smokers from Core B. Because of cost and time considerations 20 smokers with high levels of the (acrolein and crotonaldehyde) adducts measured in Project 4 Aim 1 will be selected for analysis of sub aim 2.

Statistical Analysis and Power. For sub aim 1 we will test for differences in DNA adduct levels between smokers and never smokers using parametric or (if appropriate) non-parametric single degree of freedom tests (t-test, Wilcoxon, etc.) Because it is anticipated that several hundred adducts (known + unknown) will be measured we will use a Bonferroni criteria of .05/200 = 0.00025 in order to declare that the observed differences are statistically significant. With 50 participants per group we will be able to detect (with 80 percent power) a smoking difference explaining 18 percent of variation of the adduct level, equivalent to about 1 standard deviation of the within group (smokers, never-smokers) variation of a given DNA adduct.

For sub aim 2 we analyze 20 participants per group and test for ethnic differences in adduct levels for the adducts shown to be elevated in smokers. Our main comparison is between the high risk group (native Hawaiians) and the other groups (Japanese and Whites). Assuming that 50 of the adducts are elevated in smokers we will use a p value of 0.05/50=.001 in our hypothesis testing as protection against multiple comparisons. For a given adduct we will be able to detect a difference of about 1.2 standard deviations amounting to 25 percent of variability explained by ethnicity (Native Hawaiian vs Japanese American + White).

17 Data and Biospecimen Sharing Plan

The institutions involved in this project are committed to the sharing of research data and samples, consistent with the regulations and policies of the NIH. The investigators involved in this project intend to share findings from its research through publications and presentations. Institutions and individuals wishing to access any resources or data must contact the Principal Investigator (Hatsukami). Data will be available in two formats. One

will be a summary of the data, with graphs and tables, posted as pdf files. Also, data will be available as raw individual-level data for analysis. However, this will not be available until papers are accepted for publication. Data generated by this grant will be made to outside investigators, according to NIH Guidance. Persons requesting data must do so in writing, identifying the affiliation and how the data will be used. When data are shared, there will be no limits placed on how the data will be used, and co-authorship is not required as a condition for receiving data. Users will agree, however, that the recipient must not transfer the data to other users and that the data are only to be used for research purposes. The PI will require requestors of data to sign a data and biospecimen sharing agreement that will ensure (1) Use of the data and/or samples only for research purposes unless otherwise agreed and not to identify any individual or personal information (2) Data security using appropriate technology/firewalls, (3) Destruction or return of data after data analysis and (4) Proper citation in publications or other written materials. A record of transfer of data and a copy of the dataset that was distributed will be kept by University of Minnesota.

18 Conduct of the Study

18.1 Good Clinical Practice

The study will be conducted in accordance with the appropriate regulatory requirement(s). Essential clinical documents will be maintained to demonstrate the validity of the study and the integrity of the data collected. Master files will be established at the beginning of the study, maintained for the duration of the study and retained according to the appropriate regulations.

18.2 Ethical Considerations

The study will be conducted in accordance with ethical principles founded in the Declaration of Helsinki. The UMN and UH IRBs will review all appropriate study documentation in order to safeguard the rights, safety and well-being of the patients. The protocol, consent, written information given to the patients, safety updates, annual progress reports, and any revisions to these documents will be provided to the IRBs for review by the investigators.

18.3 Informed Consent

All potential study participants will be given a copy of the IRB-approved consent to review. The investigator or designee will explain all aspects of the study in lay language and answer all questions regarding the study. If the participant decides to participate in the study, he/she will be asked to sign and date the consent document. Patients who refuse to participate or who withdraw from the study will be treated without prejudice.

References

1. Hukkanen, J., Jacob, P., III, and Benowitz, N. L. (2005) Metabolism and disposition kinetics of nicotine. Pharmacol. Rev. 57, 79-115.
2. Hecht, S. S. (2003) Tobacco carcinogens, their biomarkers, and tobacco-induced cancer. Nature Rev. Cancer. 3, 733-744.
3. Dempsey, D., Tutka, P., Jacob, P., III, Allen, F., Schoedel, K., and Tyndale, R. F. (2004) Nicotine metabolite ratio as an index of cytochrome P450 2A6 metabolic activity. Clinical Pharmacology & Therapeutics. 76, 64-72.
4. Lea, R. A., Dickson, S., and Benowitz, N. L. (2006) Within-subject variation of the salivary 3HC/COT ratio in regular daily smokers: prospects for estimating CYP2A6 enzyme activity in large-scale surveys of nicotine metabolic rate. J Anal. Toxicol. 30, 386-389.
5. Hecht, S. S. (1998) Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. Chem. Res. Toxicol. 11, 559-603.
6. Hecht, S. S. (2008) Progress and challenges in selected areas of tobacco carcinogenesis. Chem. Res. Toxicol. 21, 160-171.
7. International Agency for Research on Cancer (2007) Smokeless tobacco and tobacco-specific nitrosamines. In IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, v. 89 pp 41-583, IARC, Lyon, FR.
8. Carmella, S. G., Akerkar, S., and Hecht, S. S. (1993) Metabolites of the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in smokers' urine. Cancer Res. 53, 721-724.
9. United States Department of Health and Human Services (2010) How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Ch. 5. U.S. Department of Health and Human Services, Washington, D.C.
10. International Agency for Research on Cancer (2004) Tobacco Smoke and Involuntary Smoking. In IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, vol. 83 pp 33-1187.
11. Hecht, S. S. (2002) Human urinary carcinogen metabolites: biomarkers for investigating tobacco and cancer. Carcinogenesis. 23, 907-922.
12. Stepanov, I., Upadhyaya, P., Feuer, R., Jensen, J., Hatsukami, D. K., and Hecht, S. S. (2008) Extensive metabolic activation of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in smokers. Cancer Epidemiol. Biomarkers Prev. 17, 1764-1773.
13. Church, T. R., Anderson, K. E., Caporaso, N. E., Geisser, M. S., Le, C., Zhang, Y., Benoit, A. R., Carmella, S. G., and Hecht, S. S. (2009) A prospectively measured serum biomarker for a tobacco-specific carcinogen and lung cancer in smokers. Cancer Epidemiol. Biomarkers Prev. 18, 260-266.
14. Yuan, J. M., Koh, W. P., Murphy, S. E., Fan, Y., Wang, R., Carmella, S. G., Han, S., Wickham, K., Gao, Y. T., Yu, M. C., and Hecht, S. S. (2009) Urinary levels of tobacco-specific nitrosamine metabolites in relation to lung cancer development in two prospective cohorts of cigarette smokers. Cancer Res. 69, 2990-2995.

15. Yuan, J.-M., Gao, Y.-T., Murphy, S. E., Carmella, S. G., Wang, R., Zhong, Y., Moy, K. A., Davis, A. B., Tao, L., Chen, M., Han, S., Nelson, H. H., Yu, M. C., and Hecht, S. S. (2011) Urinary levels of cigarette smoke constituent metabolites are prospectively associated with lung cancer development in smokers. Cancer Res. 71, 6749-6757.

16. International Agency for Research on Cancer (2010) Some Non-Heterocyclic Polycyclic Aromatic Hydrocarbons and Some Related Exposures. In IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, v. 92 pp 35-818, IARC, Lyon, FR.

17. Conney, A. H., Chang, R. L., Jerina, D. M., and Wei, S. J. C. (1994) Studies on the metabolism of benzo[a]pyrene and dose-dependent differences in the mutagenic profile of its ultimate carcinogenic metabolite. Drug Metabol. Rev. 26, 125-163.

18. Thakker, D.R., Yagi, H., Levin, W., Wood, A.W., Conney, A.H., and Jerina, D.M. (1985) Polycyclic aromatic hydrocarbons: metabolic activation to ultimate carcinogens. In Bioactivation of Foreign Compounds. (Anders, M. W. Ed.), pp 177-242, Academic Press, Inc., New York.

19. Kriek, E., Rojas, M., Alexandrov, K., and Bartsch, H. (1998) Polycyclic aromatic hydrocarbon-DNA adducts in humans: relevance as biomarkers for exposure and cancer risk. Mutat. Res. 400, 215-231.

20. Zhong, Y., Carmella, S. G., Upadhyaya, P., Hochalter, J. B., Rauch, D., Oliver, A., Jensen, J., Hatsukami, D., Wang, J., Zimmerman, C., and Hecht, S. S. (2011) Immediate consequences of cigarette smoking: rapid formation of polycyclic aromatic hydrocarbon diol epoxides. Chem. Res. Toxicol. 24, 246-252.

21. Zhong, Y., Wang, J., Carmella, S. G., Hochalter, J. B., Rauch, D., Oliver, A., Jensen, J., Hatsukami, D., Upadhyaya, P., Zimmerman, C., and Hecht, S. S. (2011) Metabolism of [D10]phenanthrene to tetraols in smokers for potential lung cancer susceptibility assessment: Comparison of oral and inhalation routes of administration. J. Pharmacol. Exp. Ther. 338, 353-361.

22. Hochalter, J. B., Zhong, Y., Han, S., Carmella, S. G., and Hecht, S. S. (2011) Quantitation of a minor enantiomer of phenanthrene tetraol in human urine: correlations with levels of overall phenanthrene tetraol, benzo[a]pyrene tetraol, and 1-hydroxypyrene. Chem. Res. Toxicol. 24, 262-268.

23. Hecht, S. S., Chen, M., Yagi, H., Jerina, D. M., and Carmella, S. G. (2003) r-1,t-2,3,c-4-Tetrahydroxy-1,2,3,4-tetrahydrophenanthrene in human urine: a potential biomarker for assessing polycyclic aromatic hydrocarbon metabolic activation. Cancer Epidemiol. Biomarkers Prev. 12, 1501-1508.

24. Hecht, S. S., Chen, M., Yoder, A., Jensen, J., Hatsukami, D., Le, C., and Carmella, S. G. (2005) Longitudinal study of urinary phenanthrene metabolite ratios: effect of smoking on the diol epoxide pathway. Cancer Epidemiol. Biomarkers Prev. 14, 2969-2974.

25. Hecht, S. S., Carmella, S. G., Villalta, P. W., and Hochalter, J. B. (2010) Analysis of phenanthrene and benzo[a]pyrene tetraol enantiomers in human urine: relevance to the bay region diol epoxide hypothesis of benzo[a]pyrene carcinogenesis and to biomarker studies. Chem. Res. Toxicol. 23, 900-908.

26. Carmella, S. G., Yoder, A., and Hecht, S. S. (2006) Combined analysis of r-1,t-2,3,c-4-tetrahydroxy-1,2,3,4-tetrahydrophenanthrene and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol in smokers' plasma. Cancer Epidemiol. Biomarkers Prev. 15, 1490-1494.

27. Carmella, S. G., Chen, M., Yagi, H., Jerina, D. M., and Hecht, S. S. (2004) Analysis of phenanthrols in human urine by gas chromatography-mass spectrometry: potential use in carcinogen metabolite phenotyping. Cancer Epidemiol. Biomarkers Prev. 13, 2167-2174.

28. International Agency for Research on Cancer (2004) Tobacco Smoke and Involuntary Smoking. In IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, vol. 83 IARC, Lyon, FR.

29. International Agency for Research on Cancer (1995) Acrolein. In IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, vol. 63 pp 337-372, IARC, Lyon, FR.

30. Feng, Z., Hu, W., Hu, Y., and Tang, M.-S. (2006) Acrolein is a major cigarette-related lung cancer agent. Preferential binding at p53 mutational hotspots and inhibition of DNA repair. Proc. Natl. Acad. Sci. USA. 103, 15404-15409.

31. Zhang, S., Villalta, P. W., Wang, M., and Hecht, S. S. (2007) Detection and quantitation of acrolein-derived 1,N2-propanodeoxyguanosine adducts in human lung by liquid chromatography-electrospray ionization-tandem mass spectrometry. Chem. Res. Toxicol. 20, 565-571.

32. Zhang, S., Villalta, P. W., Wang, M., and Hecht, S. S. (2006) Analysis of crotonaldehyde- and acetaldehyde-derived 1,N2-propanodeoxyguanosine adducts in DNA from human tissues using liquid chromatography-electrsopray ionization-tandem mass spectrometry. Chem. Res. Toxicol. 19, 1386-1392.

33. Carmella, S. G., Chen, M., Zhang, Y., Zhang, S., Hatsukami, D. K., and Hecht, S. S. (2007) Quantitation of acrolein-derived 3-hydroxypropylmercapturic acid in human urine by liquid chromatography-atmospheric pressure chemical ionization-tandem mass spectrometry: effects of cigarette smoking. Chem. Res. Toxicol. 20, 986-990.

34. Scherer, G., Engl, J., Urban, M., Gilch, G., Janket, D., and Riedel, K. (2007) Relationship between machine-derived smoke yields and biomarkers in cigarette smokers in Germany. Regul. Toxicol. Pharmacol. 47, 171-183.

35. Feng, S., Roethig, H. J., Liang, Q., Kinser, R., Jin, Y., Scherer, G., Urban, M., Engl, J., and Riedel, K. (2006) Evaluation of urinary 1-hydroxypyrene, S-phenylmercapturic acid, trans,trans-muconic acid, 3-methyladenine, 3-ethyladenine, 8-hydroxy-2'-deoxyguanosine and thioethers as biomarkers of exposure to cigarette smoke. Biomarkers. 11, 28-52.

36. Scherer, G., Urban, M., Hagedorn, H. W., Feng, S., Kinser, R. D., Sarkar, M., Liang, Q., and Roethig, H. J. (2007) Determination of two mercapturic acids related to crotonaldehyde in human urine: influence of smoking. Hum. Exp. Toxicol. 26, 37-47.

37. International Agency for Research on Cancer (2012) A Review of Human Carcinogens: Arsenic, Metals, Fibres, and Dusts. In IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 100C Lyon, France.

38. Adams, S. V., Passarelli, M. N., and Newcomb, P. A. (2012) Cadmium exposure and cancer mortality in the Third National Health and Nutrition Examination Survey cohort. Occup. Environ. Med. 69, 153-156.

39. Hecht, S. S., Yuan, J.-M., and Hatsukami, D. K. (2010) Applying tobacco carcinogen and toxicant biomarkers in product regulation and cancer prevention. Chem. Res. Toxicol. 23, 1001-1008.

40. International Agency for Research on Cancer (1993) Beryllium, cadmium, mercury, and exposures in the glass manufacturing industry. In Monographs on the Carcinogenic Risk of Chemicals to Humans, v. 58 pp 41-237, IARC, Lyon, FR.

41. Brocato, J. and Costa, M. (2013) Basic mechanics of DNA methylation and the unique landscape of the DNA methylome in metal-induced carcinogenesis. Crit. Rev. Toxicol. 43, 493-514.

42. Hossain, M. B., Vahter, M., Concha, G., and Broberg, K. (2012) Low-level environmental cadmium exposure is associated with DNA hypomethylation in Argentinean women. Environ. Health Perspect. 120, 879-884.

43. Church, D. F. and Pryor, W. A. (1985) Free radical chemistry of cigarette smoke and its toxicological implications. Environ. Health Perspect. 64, 111-126.

44. Bhalla, D. K., Hirata, F., Rishi, A., and Gairola, C. G. (2009) Cigarette smoke, inflammation, and lung injury: a mechanistic perspective. Journal of Toxicology and Environmental Health, Part B. 12, 45-64.

45. United States Department of Health and Human Services (2014) The Health Consequences of Smoking: 50 Years of Progress. 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, Atlanta, GA.

46. Dedon, P.C. (2011) Oxidation and Deamination of DNA by Endogenous Sources. In Chemical Carcinogenesis. (Penning, T. M. Ed.), pp 209-225, Humana Press, New York.

47. Lee, S.H. and Blair, I.A. (2011) Lipid Peroxide–DNA Adducts. In Chemical Carcinogenesis. (Penning, T. M. Ed.), pp 227-244, Humana Press, New York.

48. Weinberg, R.A. (2007) The Biology of Cancer. Garland Science, Taylor and Francis Group, New York.

49. Frenkel, K. (1992) Carcinogen-mediated oxidant formation and oxidative DNA damage. Pharmacol Ther. 53, 127-166.

50. Ziech, D., Franco, R., Pappa, A., and Panayiotidis, M. I. (2011) Reactive oxygen species (ROS)--induced genetic and epigenetic alterations in human carcinogenesis. Mutat Res. 711, 167-173.

51. Pryor, W. A., Hales, B. J., Premovic, P. I., and Church, D. F. (1983) The radicals in cigarette tar: their nature and suggested physiological implications. Science. 220, 425-427.

52. Hecht, S. S. (1999) Tobacco smoke carcinogens and lung cancer. J. Natl. Cancer Inst. 91, 1194-1210.

53. Nair, U., Bartsch, H., and Nair, J. (2007) Lipid peroxidation-induced DNA damage in cancer-prone inflammatory diseases: A review of published adduct types and levels in humans. Free Radical Biology & Medicine. 43, 1109-1120.

54. Bartsch, H. and Nair, J. (2005) Accumulation of lipid peroxidation-derived lesions: potential lead markers for chemoprevention of inflammation driven malignancies. Mutat. Res. 591, 34-44.

55. Wang, D. and DuBois, R. N. (2013) Urinary PGE-M: a promising cancer biomarker. Cancer Prev Res (Phila). 6, 507-510.

56. Duffield-Lillico, A. J., Boyle, J. O., Zhou, X. K., Ghosh, A., Butala, G. S., Subbaramaiah, K., Newman, R. A., Morrow, J. D., Milne, G. L., and Dannenberg, A. J. (2009) Levels of prostaglandin E metabolite and leukotriene E(4) are increased in the urine of smokers: evidence that celecoxib shunts arachidonic acid into the 5-lipoxygenase pathway. Cancer Prev Res (Phila Pa). 2, 322-329.

57. Gross, N. D., Boyle, J. O., Morrow, J. D., Williams, M. K., Moskowitz, C. S., Subbaramaiah, K., Dannenberg, A. J., and Duffield-Lillico, A. J. (2005) Levels of prostaglandin E metabolite, the major urinary metabolite of prostaglandin E2, are increased in smokers. Clin. Cancer Res. 11, 6087-6093.

58. Morrow, J. D., Frei, B., Longmire, A. W., Gaziano, J. M., Lynch, S. M., Shyr, Y., Strauss, W. E., Oates, J. A., and Roberts II, L. J. (1995) Increase in circulating products of lipid peroxidation (F2-isoprostanes) in smokers. N. Engl. J. Med. 332, 1198-1203.

59. Milne, G. L., Yin, H., Hardy, K. D., Davies, S. S., and Roberts, L. J. (2011) Isoprostane generation and function. Chem Rev. 111, 5973-5996.

60. Yan, W., Byrd, G. D., and Ogden, M. W. (2007) Quantitation of isoprostane isomers in human urine from smokers and nonsmokers by LC-MS/MS. J Lipid Res. 48, 1607-1617.

61. Roethig, H. J., Munjal, S., Feng, S., Liang, Q., Sarkar, M., Walk, R. A., and Mendes, P. E. (2009) Population estimates for biomarkers of exposure to cigarette smoke in adult U.S. cigarette smokers. Nicotine Tob. Res. 11, 1216-1225.

62. Frost-Pineda, K., Liang, Q., Liu, J., Rimmer, L., Jin, Y., Feng, S., Kapur, S., Mendes, P., Roethig, H., and Sarkar, M. (2011) Biomarkers of potential harm among adult smokers and nonsmokers in the total exposure study. Nicotine Tob. Res. 13, 182-193.

63. Lonkar, P. and Dedon, P. C. (2011) Reactive species and DNA damage in chronic inflammation: Reconciling chemical mechanisms and biological fates. Int. J. Cancer. 128, 1999-2009.

64. Le Marchand, L., Derby, K. S., Murphy, S. E., Hecht, S. S., Hatsukami, D., Carmella, S. G., Tiirikainen, M., and Wang, H. (2008) Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. Cancer Res. 68, 9137-9140.

65. Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., and Fagerstrom, K. O. (1991) The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. Br. J. Addict. 86, 1119-1127.

66. Smith, S. S., Piper, M. E., Bolt, D. M., Fiore, M. C., Wetter, D.W., Cinciripini, P. M., and Baker, T. B. (2010) Development of the Brief Wisconsin Inventory of Smoking Dependence Motives. Nicotine Tob. Res. 12(5), 489-499.

67. Task Force on Recommended Alcohol Questions. (2003) National Council on Alcohol Abuse and Alcoholism recommended sets of alcohol consumption questions. National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, https://www.niaaa.nih.gov/research/guidelines-and-resources/recommended-alcohol-questions.

68. Selzer, M. L. (1971) The Michigan Alcohol Screening Test: the quest for a new diagnostic instrument. Am. J. Psychiatry 127(12), 1653-1658.

69. Nondahl, D. M., Cruickshanks, K. J., and Schubert, C. R. (2005) A questionnaire for assessing environmental tobacco smoke exposure. Environ. Res. 97, 76-82.

70. Routt, R. and Roberts, J. R. (1999) Environmental and occupational history. In Recognition and Management of Pesticide Poisonings. (Routt, R. and Roberts, J. R., Eds.), U.S. Environmental Protection Agency, Washington, D.C.

Appendix A: Eligibility Checklist:

Observation Study

Participant initials ☐☐☐☐ Patient ID ☐☐☐☐

Eligibility Checklist – page 1 of 2

INCLUSION CRITERIA

A "NO" response to any of the following disqualifies the participant from study entry.

		Yes	No
1.	One of the three targeted ethnic groups: a. Japanese American - two parents of Japanese descent. ☐ b. Non-Hispanic Whites -- two parents of non-Hispanic white descent. ☐ c. Native Hawaiians will include individuals with at least one parent of Hawaiian descent; ☐	☐	☐
2.	Smokes 5 cigarettes per day in the past three months;	☐	☐
3.	≥ 21 years of age	☐	☐
4.	Consumes 14 or fewer drinks of alcohol per week;	☐	☐
5.	Generally good health (determined by review of medical history)	☐	☐
6.	Able to provide written voluntary consent before performance of any study related procedure.	☐	☐

Eligibility Checklist – page 2 of 2

EXCLUSION CRITERION

A "YES" response to the following disqualifies the participant from study entry.

		Yes	No
7.	Current use of other nicotine containing products for > 4 times per month (and no use of any nicotine-containing products except cigarettes for 2 weeks prior to their study visits);	☐	☐
8.	Acute or uncontrolled medical or psychiatric conditions;	☐	☐
9.	Currently taking any medications that affect relevant metabolic enzymes or anti-inflammatory medications such as ibuprofen (this will be reviewed by study investigators on a case-by-case basis);	☐	☐
10.	Active infection (e.g., influenza, cold, respiratory infection, sinus infection) at the time of the visit;	☐	☐
11.	Pregnant or nursing or planning on becoming pregnant during the study;.	☐	☐
12.	Unable to read and understand English.	☐	☐

Date consent form signed by participant and research staff:_____

Having obtained consent and reviewed each of the inclusion/exclusion criteria, I verify that this participant is:

☐ Eligible ☐ Ineligible Date registered_____

_____ _____

Signature of person verifying eligibility Date