Supporting Statement B For:

STUDY TO ESTIMATE RADIATION DOSES AND CANCER RISKS FROM RADIOACTIVE FALLOUT FROM THE TRINITY NUCLEAR TEST (NCI)

Date: TBD

Steven L. Simon
Radiation Epidemiology Branch
Division of Cancer Epidemiology and Genetics (DCEG)
National Cancer Institute
National Institutes of Health

9609 Medical Center Drive
Rockville, MD 20850

Telephone: 240-276-7371
E-mail: ssimon@mail.nih.gov

Table of Contents

LIST OF ATTACHMENTS

B. STATISTICAL METHODS

B.1 Respondent Universe and Sampling Methods

Sampling Frame

The respondent universe consists of ethnic groups in New Mexico for which little data relevant to 1945 is available, in particular, for the Native American and the Hispanic groups, as well as non-Hispanic whites.

Potential participant subjects may include any of the following: community elders, tribal members, and people with firsthand knowledge of food practices in New Mexico in the 1940s or 1950s. Other key informants, such as tribal health representatives, *promotores de salud*, historians, representatives from academia, and advocates who can provide firsthand knowledge and/or academic expertise about the communities that can provide additional sources of information will be collaborated with. The population (2010 U.S. Census) of New Mexico of persons ≥ 70 is over 184,000. In New Mexico, 46% of the population is Hispanic/Latino, 41% of the population is non-Hispanic white, and 9% Native American.

Inclusion criteria

To be eligible to participate in focus groups or interviews, individuals must self-identify as a New Mexico tribal community member, Hispanic/Latino or non-Hispanic white, and have firsthand knowledge about lifestyle and dietary patterns in New Mexico during the 1940s or 1950s. According to the 1940 census, New Mexico had a population size of 531,818 with 275,427 whites (52%), 34,510 (6%) self-identifying as Native American and 221,881 (42%) Hispanic/Latino (of any race). The focus group participants will include non-Hispanic whites, Hispanics and Native Americans who are currently ≥ 70 years old and in the 1940s resided in communities in or near the fallout region in NM (**Attachment 2**) including Native American pueblos and tribes. Participants will be identified from Native American experts and community liaisons. The focus groups will be conducted in different ecoregions or urban/rural settings among the various ethnic groups.

Sample Size

The study team plans to recruit up to 210 elders to participate in the focus groups and interviews. Fliers will be distributed in local communities in both English and Spanish (**Attachment 4)**, and subjects will be screened for eligibility, in either English or Spanish, (**Attachment 5**) and provided with a letter, in either English or Spanish, about the study prior to the conduct of any interviews (**Attachment 6)**. It is anticipated that 50% of individuals screened will participate in either a focus group or interview. The study team intends to conduct 21 focus groups with up to 8 participants in each group. For comparison purposes, eight focus groups on the lifestyle and dietary patterns of non-Hispanic whites will be included. In addition, we will conduct up to 42 key informant interviews. The table below describes the number of data values that will be collected from the focus group participants.

Race/ethnicity	**Native American**			**Hispanic/Latino**				**Non-Hispanic White**		
Number of focus groups	5			8				8		
Residence	Mountain	Desert plains		Urban north	Urban south	Small town Rural		Urban north	Urban south	Small town /Rural
Number of focus groups	2	3		2	2	4		2	2	4
Participants[1]	16	24		16	16	32		16	16	32
Response cells about women[2]	96	144		96	96	192		96	96	192
Response cells about men[3]	80	120		80	80	160		80	80	160
Response cells (total)	176	264		176	176	352		176	176	352

[1] *Assumes 8 members of each focus group*

[2] *Assumes 5 age groups (infant, 1-4, 5-10, 11-15, 16+) plus pregnant or breastfeeding women*

[3] *Assumes 5 age groups (infant, 1-4, 5-10, 11-15, 16+)*

B.2 Procedures for the Collection of Information

Participant Identification and Recruitment

Potential focus group members and key informants may include any of the following: community elders, tribal members, and people with firsthand knowledge of food practices in New Mexico in the 1940s or 1950s. To be eligible, individuals must self-identify as a New Mexico tribal community member, Hispanic/Latino or non-Hispanic white in order to participate in an interview, and have firsthand knowledge about lifestyle and dietary patterns in New Mexico during the 1940s or 1950s. According to the 1940 census, New Mexico had a population size of 531,818 with 275,427 whites (52%), 34,510 (6%) self-identifying as Native American and 221,881 (42%) Hispanic/Latino (of any race)(U.S. Census Bureau, 1940).

We plan to recruit up to 210 elders to participate in the focus groups and interviews. A university contractor in New Mexico will collaborate with community groups, e.g., *Las Mujeres Hablan* (a network of women-led organizations in NM interested in health and nuclear issues), community health representatives, *promotoras de salud* and other community leaders engaged with the NCI team to identify and recruit potential participants for the focus groups. Subjects will be screened for eligibility (**Attachment 5**) and provided with a letter about the study prior to the conduct of any interviews (**Attachment 6**). The scheduling and coordination of the focus groups will be under the direction of the university contractor. Once all IRB and OMB approvals to conduct interviews are in place and appropriate preliminary meetings have taken place with community leaders and potential participants, the focus groups will be scheduled. If the intended participants do not attend the focus group at the designated time, any eligible persons arriving will be consented and interviewed. If only one person attends, he or she will be interviewed as a key informant rather than as focus group member. If a participant would rather not be interviewed in a group, he or she will be interviewed as a key informant.

We will target recruitment efforts for people who resided in geographic areas that lie within or on the outskirts of the primary fallout pattern from the Trinity test (**Attachment 2)**.

Approval to conduct focus groups and interviews in the noted communities will be obtained from appropriate community leaders prior to recruiting study participants. IRB approval has been obtained from the Albuquerque Area Indian Health Board for the focus groups and key informant interviews, and tribal resolutions from two Tribal Nations have been obtained. Tribal resolutions will be obtained for any Tribal Nation where we will conduct focus groups. NCI REB support services contractor, Social & Scientific Systems (SSS), will established a subcontract with a university researcher in New Mexico in order to allow us to collaborate and conducted the interviews and focus groups. The university contractor will be responsible for the following tasks:

- Requesting and securing required Tribal Resolutions
- Conduct study outreach in local communities
- Recruiting participants
- Issuing compensation
- Securing interview meeting locations
- Transportation
- Providing food and refreshments
- Co-moderating during focus groups/interviews
- Hiring translators and interpreters as needed

Data Collection Procedures

The main purpose of the data collection is to quantify the range of dietary intakes for different categories of age, sex, ethnic group, and ecoregion. In addition we seek information about other radiation-related exposure pathways such as building materials of houses and schools, and time spent outside by sex and age (all relevant to assessing external dose). Quantitative and qualitative data recalled about the individuals' family will be collected through the focus group sessions while more general information about the community will be collected through the Key Informant interviews.

This approach, to ascertain prior exposure to nuclear radiation is a qualitative research technique that relies heavily on probing questions by the interviewer to elicit thoughts, memories, and interpretations by subjects, and can be related to key events in their past. The procedure is often used in the testing and development of survey questionnaires (Willis, 2005), but can also be used more flexibly to reconstruct events in one's past (Belli & Callegaro, 2009). As such, the technique is well-suited for the collection of oral histories of the type to be collected within the current study. In brief, cognitive interviewing involves the use of probe questions that lead the subject to elaborate on the topic under discussion. Given that human memory, especially of longdistant events, tends to be reconstructive, probing takes advantage of this by helping the individual to reinstate the context surrounding the memories, and to rely on recalled information as cues to elicit further memories.

Focus group moderator guide

The Trinity focus group moderator guide was designed using the REB study of fallout exposure in Kazakhstan (Land et al., 2015; Land et al., 2008; Schwerin et al., 2010) as guide. It includes detailed probes in the form of open-ended questions, and data collection sheets carefully designed to elicit the data sought. Prior to the focus group discussions, participants will be greeted, consented (**Attachments 8 and 14**), and complete a table about their family members. They will then do a card sorting exercise related to the frequency food items that were consumed in their families. The pre-focus group guide is available in **Attachment 8**. The moderator focus group guide (**Attachment 9**) is divided into the following sections: Greetings, Logistics, Introductions, Daily Life at the Time of the Test, and Food and Drink (frequency and quantities consumed). A specific amount of time is allowed for each section to ensure that all sections are covered during the two-hour time given to each focus group.

Focus group exercise

After the consenting process and introductions, the moderator will begin with discussion of the time period, living situations, number of family members, etc. The moderator will direct the discussions and allow for members to deliberate as needed on various topics.

Card sorting

The focus group interview will begin with a 10-minute card sorting exercise. Card sorting is a useful mixed-method tool which allows for quantitative and qualitative mapping of concepts from the participants' point of view. From the list of foods identified in the pilot study, those foods that are important sources of exposure to radionuclides were selected for a card sorting exercise. Cards will be printed with the image and name of each food of interest. Each participant will be asked individually to sort the cards into two piles: those foods consumed by his/her community in 1945 (or 1940s and 1950s), and those foods that were not. This exercise aims to determine the foods consumed by specific communities. Sorting through the images helps to familiarize the participant with the food items that we are most interested in quantifying during the focus group interview. From the many types of foods that were consumed, participants will then be asked to sort the foods according to food groups and then rank the foods in order from most to least frequently consumed. The sorted and ranked cards will be stapled together in their piles and will be later recorded and collated with the rest of the information gathered in the wall chart exercise (described next) of the focus group.

Wall Chart

Questions will be asked about consumption patterns of the individual's family members; quantification will be made for each food item, for each age group, and for each gender. The information will be recorded on a wall chart by the co-moderator as well as by the note takers. The types of food will be listed across the top of the table and the age groups will be listed in the rows of the table (**Attachment 9**). Each participant will be asked to estimate the amount of each food consumed by each age group using samples of typical serving ware, which were identified during the pilot study. This method was previously used by members of the

Trinity Team and was proven effective in the reconstruction of diet from the distant past among an elderly Kazakhstan population (Schwerin et al., 2010).

If permission is given by all participants, there will be an audio recording made of the session. Translators/interpreters with experience in the study populations and who are certified to translate and interpret will be present when needed. Participants will not be identified by full name in any recorded materials. Audio recordings will be destroyed at the conclusion of the study. At the end of the session, participants will be transported home as needed and compensation will be mailed to them. Participants can leave the session at any time and will still receive their compensation and transportation.

Following each focus group session, there will be a debriefing with the focus group moderators, NCI team members, and any consultants on site. Any items that may need further clarification, elaboration or verification will be incorporated into the subsequent focus group written record. Any modification of procedures or further clarification of a topic will be incorporated into subsequent focus groups. Upon completion of focus group sessions in each geographic region, the research team will gather to review the notes taken by the co-moderator and a second member of the research team. During that time, lists of recurring themes will be generated, typical quotes from participants will be noted, and terms, phrases and findings that have language or cultural significance and may not be conveyed well when translated, will be discussed.

Upon completion of the data collection phase, we will use transcripts, notes and charts to prepare a report of the focus groups. These notes and supplementary material will be used to help assess the nature and extent of uncertainty of what was reported during the focus group. Preparation of the focus group reports will be assisted through consultation with SSS (Social & Scientific Systems). Quantitative estimates of food intake and other lifestyle descriptions (e.g., typical time outdoors) will be synthesized into tables for development of the models of lifestyle and diet to be used in the exposure assessment.

In-depth interviews with Key Informants

For each diet or lifestyle topic, there is a series of questions. The study team has designed an in-depth key informant interview guide (**Attachment 7**), including detailed open-ended questions to help stimulate participant memory. One team member will conduct each individual interview, while another will be responsible for recording the information. The interviewers will encourage the subject to provide an estimate or answer to all of the questions posed.

The topics of the questions cover the roles of women and men within their social structure, specifically with regard to the production and consumption of food during the time of the Trinity test as well as lifestyle habits of community members. The interview guide is divided into sections addressing dairy, meat, and vegetable consumption, source, and storage of milk/milk products. It also focuses on building materials of houses and schools, availability of milk and milk products, home remedies, time outdoors, and grazing practices. There is a specific amount of time allotted to each section to ensure that all sections are covered during each in-depth interview. The guide does not elicit information about quantities of dietary constituents consumed. This in-depth interview guide was pilot tested and proved to be an effective tool during the Phase 1 study.

Prepare focus group data for analysis

During the focus groups and interviews, the study team (composed of two NCI moderators, one with extensive experience working with tribal communities and one with Hispanic communities, and up to two observers) will take notes and record numeric data (e.g., age-specific estimates of the quantity of cow milk consumed) onto the moderator guide (**Attachment 9**).

Detailed transcripts of the interview and provision of spreadsheets from the wall charts will be prepared by the REB support services contractor, SSS, following the completion of the fieldwork. Trinity study investigators will abstract the results of the focus/groups interviews with regard to key radiation exposure pathways and create a database to be used in dose reconstruction efforts. The transcripts and data collection sheets will summarize the interview data and prepare them for calculating summary statistics as a basic model of the Native American, Hispanic, and non-Hispanic White diets with regard to key radiation exposure pathways. From the transcripts and data collection sheets, variables will be created and used for summarizing the results of the interviews.

Data quality measures will be a major component of the interviewer approach. The study team will use focus group and interviewing guides to elicit information in which all aspects of the study protocol are standardized and described in detail. Each section of the guide will have a specific time allotted so that all sections are covered during each focus group session or in-depth interview. As stated above, the note-takers will collect responses on data collection sheets thus improving the consistency of questions across interviewees. For focus groups, the information collected on data sheets will be supplemented by data collected on a large wall chart. The audio recordings of the interview session will be used to generate and check transcripts; this will improve the reliability of the data.

In an effort to assess the quality of the focus group data obtained in the proposed study, we have included 8 focus groups comprised of Non-Hispanic whites. For efficiency, they will be in the same locations as planned focus groups for Hispanics within the fallout zones. These focus groups will be conducted in the following ecoregions; a small town in a mountainous region in the south, an urban area in a mountainous region in the north, an urban area in a plains region in the north, and a rural area in a plains region in the central area.

The data obtained in the focus groups will be compared with published data on “Estimation of 1945 to 1957 Food Consumption” prepared for the Hanford Technical Steering Panel and the Centers for Disease Control and Prevention (Anderson et al., 1993). In that compendium, data were taken from a 1977-78 USDA National Food Consumption survey and estimated intakes back to particular years based on food disappearance data as well as home production for rural areas. While their focus was on Oregon and Washington State, they evaluated differences by region for intakes of key foods, including milk, and found no difference with the region that included New Mexico but significant differences with other regions. Sensitivity of their backcasting method was tested against local and national surveys that were conducted between 1965 and 1969. They found better agreement of their backcasting method for particular years of the local surveys than with a USDA survey conducted in 1965-66. Given their careful methods and agreement with smaller surveys, the assumption will be made that their method will also be appropriate for the 1945 time period. The data are available by urban or rural status, season, age and sex groups that match the groupings in our focus group guide.

We will compare intakes of milk, spinach, eggs and lettuce, each grouped into 4 categories, stratified by age and sex, estimated from the focus groups with those in the published data using chi-square tests. We will consider $p<0.01$ a statistically significant difference to account for multiple testing. If there is systematic under or over-reporting for all four foods we will consider using that information to correct the results from the focus groups (understanding that it is based on very small sample sizes).

Overall, if the ranges for a given item are consistent between communities/interviewees, age-specific summary tables will be created for reported milk and dairy consumption, other food consumption (e.g. green leafy vegetables, grain, fruit, and meat), water consumption and source, building construction materials, and time spent outdoors. In addition to summary statistics (e.g. mean), these tables will contain measures of variability (e.g. standard deviation, minimum, maximum).

Estimation Procedures

Radiation dose

When estimating radiation doses, it is necessary to distinguish between *external* and *internal* sources of radiation. *External* exposure occurs when radiation from a source outside of the body penetrates the body and is absorbed within it, while *internal* irradiation results from exposures to radionuclides that are within the body, i.e., radioactive materials that have been taken into the body through ingestion or inhalation.

For estimating both external and internal radiation exposure from fallout, the most essential input data are historical measurements of exposure-rate from radioactive fallout deposited across New Mexico at a specified time (relative to the detonation time). Our dose assessment is to be based on the analysis of fallout exposure-rates (Quinn, 1987) which extended from the detonation site towards the northeast supplemented with an extension by Cederwall and Peterson (Cederwall & Peterson, 1990) to the northern border of NM based on meteorological analysis.

The deposition density (radioactivity of each nuclide deposited per unit area of ground) of individual fallout radionuclides will be estimated from published factors (Hicks, 1981; Hicks, 1982; Hicks, 1985; Hicks, 1990) and the exposure-rates decay corrected to time of deposition. Deposition density on plant materials, more correctly called *interception*, will be estimated for leaves of vegetables, plants, and fruits that contribute to dietary intake. Interception calculations will use historical data on precipitation in New Mexico in the 48 hours following the Trinity test. Plants known to have been consumed by grazing animals and New Mexico populations are listed in Table 3 of **Attachment 3**. Based on considerations of standing biomass and leaf area, plants will be grouped by similarities such that a simplified interception modeling strategy can be developed.

Using a series of radiation exposure models that have been developed for other assessments of fallout exposures but adapted for the Trinity test, information from estimated deposition density on the ground and of interception onto plants will be combined with data on lifestyle and dietary patterns to estimate internal and external organ doses received by the population of New Mexico. This strategy was used in the study of exposures in Kazakhstan (Land et al., 2015; Land et al., 2008) and in the Marshall Islands risk projection (Land et al., 2010).

External irradiation after Trinity was dominated by gamma radiation from radioactive decay of fallout deposited on the ground. The model for external dose calculation is presented in detail by

Simon and colleagues (Simon et al., 2006). They estimated the external dose received in villages near the site of the first Soviet nuclear test in Kazakhstan, which was a nuclear device nearly identical to Trinity. Individuals receive higher doses when they are outside than when they are indoors because the gamma radiation is reduced by the shielding provided by building materials. The magnitude of the shielding effect depends on the materials (e.g., wood, adobe, or stone) and design of the building. Typically, thick adobe (mud) walls provide greater shielding from outdoor radiation than do older-style wood frame houses. In the NCI's 2008 report to Senator Bingaman, it was assumed that construction materials used for houses and buildings were those typical to NM (National Cancer Institute, 2008a, 2008b). Moreover, estimates of the typical numbers of hours spent outdoors by different age groups were used in dose calculations. None of those values have yet been corroborated by field data collection.

Internal irradiation results from exposures to radionuclides deposited in the body, mainly via ingestion of contaminated foods or water, but also by inhalation of contaminated air. In addition to fresh milk, animal tissues and organs as well as all plants can be contaminated with radioactive iodines, especially Iodine-131 (I-131 or ^{131}I) and Iodine-133 (I-133 or ^{133}I). Generally, fresh milk, leafy vegetables and possibly animal organs (e.g., thyroid glands-if consumed) are the most important contributors to the dose to the thyroid gland. The thyroid is the most highly exposed of all organs. Mother's breast milk can also be contaminated with radioiodines if the woman consumed locally produced dairy products. Factors to quantitatively estimate the concentration of radioiodine in mother's breast milk are available (Simon et al., 2002). A list of foods consumed is presented in **Attachment 3**.

For organs and tissues other than the thyroid gland for which the radiation dose will be assessed, including stomach, colon, and red bone marrow, fallout radionuclides other than radioiodines typically dominate the dose. For example, for the case of Marshall Islanders exposed to radioactive fallout, Simon et al. (Simon et al., 2010) calculated that the internal dose to the colon was dominated by ^{239}Np, ^{132}Te, ^{140}Ba, ^{99}Mo, ^{140}La, while exposure of the red bone marrow was dominated by ^{132}Te, ^{89}Sr, ^{140}Ba, ^{99}Mo, and ^{239}Np, and exposure of the stomach was dominated by ^{239}Np, ^{132}Te, ^{132}I, ^{133}I, ^{99}Mo, and ^{140}La.

Internal dose may result from consumption of foods other than dairy foods though milk products typically contain the highest concentration of radioiodines. Drinking water can contain radioactive contamination though the magnitude depends on the source of the water, physical factors that would affect dilution (e.g., stream flow) and the degree to which the water was exposed to the open atmosphere. Open water catchments would be more susceptible to contamination than well water. Consumption of leafy vegetables, herbs, and medicinal plants, whose leaves can collect and retain fallout particles, can also result in significant intakes by man. In addition, meat and bread or other flour-based foods can contain radioactive contamination though they are more susceptible to contamination with longer lived radionuclides, e.g., ^{137}Cs.

In addition to the estimation of radionuclide intakes via dairy foods, the dose assessment will incorporate the estimate of radionuclide intakes via drinking water, leafy vegetables, and animal flesh, in particular, organs such as the thyroid of small mammals that may have been consumed as part of the diets of tribal communities (Harper et al., 2002). The dose assessment will also account for special sources of radionuclide intake, such as breastfeeding (Simon et al., 2002) and inadvertent and intentional ingestion of soil (Maxwell & Anspaugh, 2011; National Council on

Radiation Protection and Measurements, 1999; Simon, 1998), a pathway of concern in dry, dusty conditions, e.g., found in the desert environment of New Mexico.

Pathway dose modeling and established dosimetry models for fallout (Beck et al., 2006; Kleinhans, 2010; National Cancer Institute, 1997) will use air, food, and water intake rates as derived from published data for white populations (Kleinhans, 2010; National Cancer Institute, 1997) combined with focus group findings and the data obtained from interviews with tribal community members, Hispanics/Latinos, and non-Hispanic whites combined with the data obtained from literature, Phase 1 field notes to estimate typical age-specific radiation organ doses per unit of radioactivity deposited on the ground for each of the three ethnic groups. Using those values, typical ranges of radiation doses to residents can be estimated within each county by combining the estimates of dose per unit deposition with the radioactivity deposition densities that will be estimated for each county precinct of New Mexico. Finally, the average dose with each county for groups defined by age group, gender, and ethnicity will be derived based on weighting by the population size in each area of the county with differing fallout deposition

Several factors and calculations will be updated, compared to the 2008 calculations, in order to improve the previous exposure assessment. These factors include:

- Based on list of plants from Phase 1 interviews, a review of fallout interception on plants based on plant characteristics and their availability within each ecozone,
- Geographic distribution of the radionuclide contamination on pasture grass following the detonation estimated from the aforementioned exposure-rate monitoring data,
- Occurrence of precipitation, and the local rainfall rate (derived from historical meteorological records),
- Radionuclide concentration in cow and goat's milk (with emphasis on radioactive iodines) estimated by typical amounts of pasture grass consumed by the animal per day and the location-specific contamination of the grass,
- Typical values and ranges of daily consumption rates for food items of interest for representative sex and age groups (adults (16+), 11-15-years old, 5-10-years old, 1-4-years old, infant to 1- year old, breastfeeding mothers, pregnant women),
- Age-specific organ dose coefficients corresponding to intake of each radionuclide as derived and recommended by the International Commission on Radiological Protection (*98*).

Diet and Lifestyle Models

The model diets, as noted earlier in this Protocol, are intended to be age-, ethnic group-, ecoregion- (and possibly gender-) specific, and to include descriptions of the foods regularly consumed in summer months in the mid-1940s as well as quantitative estimates of the frequency and quantity (measured in mass) of each. While such an undertaking would be difficult to impossible to achieve for individuals, a risk projection only requires that diets and doses be assessed for representative (typical) persons in each strata (age, gender, ethnicity, ecoregion). The data from the focus groups to be conducted, supplemented with data obtained from key informant interviews, will be used to assess mean values (i.e., representative) and ranges of the amounts of foods consumed. The strategy of using focus group and interview information to develop model diets was used in the REB study of thyroid disease in Kazakhstan (Land et al., 2015) - a study of exposure to radioactive fallout from Soviet nuclear tests which included the Soviet test of a device that replicated Trinity. Details on the data collected and the model diets constructed were published by Schwerin et al, 2010 (Schwerin et al., 2010).

Lifestyle models emphasize exposure-related attributes that differ from dietary intake. The primary variables of interest are the materials used for home construction and amount of time spent in- and outdoors during summer months. Other variables that may be identified in focus groups and interviews will be incorporated as dictated by the availability of data.

Models will be derived based on literature, Phase 1, and Phase 2 data.

In addition to point estimates of each parameter, a range of values will be generated from the data in order to quantify the uncertainty.

Dose and Uncertainty Assessment

Methods of assessing exposures are described above. In this section, uncertainty of dose is addressed.

Uncertainty of diets and related radiation doses will be assessed, though in less detail than for the Kazakhstan analytic epidemiologic study. In this study, the data collected will be used to estimate plausible ranges of diet and lifestyle and to propose probability density functions to describe variation within each county and stratum (where strata are defined by ethnicity, age, and gender). Well-developed uncertainty analysis techniques based on Monte Carlo sampling (numerical simulation) will be used to assess uncertainty in dose and in risk.

In short, the organ dose to any person is a function of physical factors (*PF* in eq. 1 below) related to fallout deposition and concentration in foodstuffs, human factors (*HF* in eq. 1, e.g., diet, dwelling type, time spent in and outdoors), and dose conversion factors (*DF* in eq. 1) which convert the radioactivity intake (for internal exposure) and the air exposure rate (for external exposure) to radiation dose:

$$D = f(PF, HF, DF) \quad (1)$$

Physical factors are numerous and include the parameters of fallout deposition that are partially summarized in the fallout map provided in Appendix A but also include fallout transit time, factors to describe interception by plants, weathering effects, and that quantitatively describe uptake and retention in foodstuffs. A detailed accounting of all physical factors was conducted in the NCI study of thyroid disease in Kazakhstan (Land et al., 2015) but is also employed in the NCI fallout dose calculator that is now being updated (http://dceg.cancer.gov/tools/public-data/risk-calculator-offline), studies of Chernobyl fallout (Drozdovitch et al., 2015), and studies of Nevada Test Site fallout (Beck et al., 2006; National Cancer Institute, 1997; Till et al., 1995).

Human factors in this context are diet (daily intakes of each food type by strata) and lifestyle parameters, in particular, type of residential construction, and time spent in and outdoors in summer months.

In the dose assessment strategy for this study, three doses within each county and strata will be estimated: a high value, a mean value, and a low value. These doses will effectively encompass the plausible range of true dose for anyone who self-identifies in a particular stratum. For this purpose, the minimum, mean, or maximum dose will use the minimum, mean, or maximum value of fallout (from the map in **Attachment 2**), respectively, in county. The numerical values of the other physical factors needed to compute dose will be median values derived from literature-based probability density function that are summarized in numerous publications. Similarly, low, mean, or maximum of the human factors (as derived from the focus groups) for each strata will be used to estimate the low, mean, and maximum dose per county.

The combination of PF, HF, and DF factors needed to estimate the point estimate and the uncertainty of the low, mean, and high doses (by strata within each county) will be computed from a Monte Carlo sampling of the relevant parameter distributions.

Risk Projection

Overview

The analytic approach for risk projection in the Trinity study will be similar to that used in the cancer risk projection conducted for the population of the Marshall Islands who were also exposed to radioactive fallout from nuclear testing (Land et al., 2010) with some specific differences. We will estimate external and internal radiation dose to the four tissues of interest. The first difference is that we will follow the recommendations of the National Academy of Sciences in their Biological Effects of Ionizing Radiation (BEIR VII) report which was released shortly prior to the Marshall Islands work (Land et al., 2010). Second, these doses will be applied to baseline cancer rates, lifetables, and published sex- and organ-specific risk coefficients to derive the corresponding excess cancer cases per 1,000 persons in each group specified by certain strata such as age at exposure, sex, and ethnicity. Cancer risk estimates per 1,000 will be presented in terms of the number of cancers by organ site that already likely occurred and that are projected to occur among non-Hispanic whites, Hispanics/Latinos, and Native American populations in New Mexico as a consequence of exposure to fallout from the Trinity test.

Study population

We have obtained several sources of underlying population counts for these analyses: the 1930, 1940, and 1950 U.S. census reports which contain population counts by county, in addition to the Handbook of North American Indians (Ortiz, 1979; Sturtevant & Ortiz, 1983), which contains more detailed information of population distributions of various tribal communities in New Mexico. In brief, the total population of New Mexico in 1940 was reported to amount to 531,818 people, including over 270,431 non-Hispanic whites, 221,881 Hispanics (of any race), and 34,510 Native Americans (U.S. Census Bureau, 1940; Ortiz, 1979).

While the number of non-Hispanic whites and Hispanic whites combined is available by county through the 1940 census, the number of people in each of the two ethnic groups is only available for the state of NM as a whole through a 5% sample. Several sources of U.S. census data will be used to estimate the number of Hispanic whites separately from non-Hispanic whites by county. The 1930 census contained a "race" category for "Mexican" which represented birth place in Mexico or parents' birth place in Mexico. This definition would suggest that Hispanic whites represent approximately 15% of the total NM population. The 1940 census, which was based on a 5% sample, indicates that Hispanic whites (using the definition of Spanish mother tongue) represent 41.7% of the entire population of NM. We will use the distributions of "Mexican race" across counties based on the 1930 census scaled to the more inclusive 41.7%. We will set the number of Hispanic whites to equal the total population when the estimated number of Hispanic whites exceeds the known total population for that county. To obtain the total population alive in 1945 during the time of the test, we will interpolate the 1940 and 1950 census population estimates.

Baseline cancer rates

Incidence rates for specific cancer sites vary from population to population, as well as over calendar time, across ages/birth cohorts, sex, and race/ethnicity groups within a population. New Mexico state cancer incidence rates prior to the establishment of the NM SEER registry in 1973 are limited (Jemal et al., 2010), but we are currently working with the head of the NM tumor registry to obtain cancer incidence data by age as early as 1969 for the state of NM (Wiggins, 2015). Cancer incidence rates for Hispanic whites became available in 1981 through SEER. Native American cancer incidence rates are available through the "Other" category in the NM registry and through the Arizona Indian Registry, which was also established in 1973. Cancer rates have been historically lower in Hispanic whites in NM than in non-Hispanic whites. In the absence of state data, national data from the first three major cancer incidence surveys conducted in the United States have been obtained. These are the Ten Cities Surveys of 1937 and 1947 (also referred to as the First and Second National Cancer Surveys), and the Iowa study of 1950 (Dorn, 1944; National Institutes of Health, 1954; National Institutes of Health, 1955). The areas surveyed included about 10 percent of the U.S. population and was representative of the geographic distribution of Northern, Southern, and Western cities with populations greater than 100,000, but was not entirely demographically representative of the U.S. population. Denver, Colorado was the city closest to New Mexico. From 1950 to 1957, the End Results in Cancer program provides cancer incidence by age from 4 central registries and 9 hospital registries across the United States (Crittenden et al., 1961). We have obtained these documents providing cancer incidence to either use directly or triangulate with the preferred method of estimating cancer incidence.

Mortality rates are available for the whole of the United States from 1930 through the National Center for Health Statistics. Cancer mortality in New Mexico for non-Hispanic whites, Hispanic whites, and Native Americans have been estimated for the period of 1958-1982 (Becker, 1993). In the absence of cancer incidence data for NM by race/ethnicity, for more lethal cancer sites we may use mortality rates to estimate incidence by calculating the ratio of incidence to mortality during a period when both measures were available and extrapolating back in time to the calendar period of interest. To improve these estimates, we may incorporate a lag-time, i.e., time from cancer diagnosis to death available from hospital based registries in New Mexico.

To the degree possible, the limitations of the limited baseline cancer rate data will be addressed by uncertainty analysis techniques as follows. We will linearly extrapolate rates from 1969 to 1950 by fitting Poisson regression models adjusted for sex and racial/ethnic group. Sources of uncertainty in our risk projection model will include those inherent in the calculation of baseline cancer rates and due to the extrapolation method as well as the estimates of average organ dose within each stratum. To obtain confidence intervals for excess cancer case estimates, we will perform a parametric bootstrap procedure for each cancer site. In that method, we will generate bootstrap case counts in each demographic stratum by drawing from a Poisson distribution whose mean is the count predicted by the risk projection model. We will then re-fit the Poisson model to the resulting case counts and use the re-fitted models to predict cancer rates. We will repeat this procedure 1000 times, and then compute the variance of the prediction based on the 1000 predicted values. This approach will provide variance estimates that account for all sources of variation (under some mild assumptions). We will provide a detailed table of the baseline cancer rates that were used for our calculations in a final manuscript. Each factor used in the calculation of the excess number of cancers per 1,000 people has an associated uncertainty, including radiation doses, parameter values of dose-response models, and other factors used for the risk projections. The uncertainty of each component will be characterized by probability distribution functions and the error propagated by Monte Carlo methods as in previous REB studies on exposures and risks in the Marshall Islands and in Kazakhstan and as described in previous paragraphs.

Radiation risk models

Weighted BEIR VII dose-response models will be used for estimating the excess relative risk and excess absolute risk per unit dose of radiation (National Research Council, 2006). Most of the radiation dose-response coefficients used for risk projection are based on the analyses of the Japanese atomic bomb survivors (Preston et al., 2007), which are considered the gold standard in radiation risk assessment. However, unlike the acute, direct external radiation exposures experienced by persons exposed to the Hiroshima and Nagasaki atomic bombings, the exposures associated with the Trinity fallout event will be considered to be continuous. Latency periods of different lengths will be considered, depending on the cancer site using an approach developed by the National Institutes of Health (National Cancer Institute, NCI-CDC Working Group to Revise the NIH Radioepidemiological Tables, 2003; National Institutes of Health, 1985).

Population distribution, race/ethnicity-, ecoregion-, organ-, sex-, age-specific doses, and organ- and sex-specific risk coefficients will be used to project the corresponding cancer risk estimates for the different population subgroups considered.

Uncertainty

Sources of uncertainty in our risk projection model will include those inherent in the calculation of baseline cancer rates and due to the extrapolation method as well as the estimates of average organ dose within each subgroup (e.g., age at exposure, gender, and ethnicity). To obtain confidence intervals for excess cancer case estimates, we will perform a parametric bootstrap procedure for each cancer site (Robbins et al., 2015). We will generate bootstrap case counts in each demographic stratum by drawing from a Poisson distribution whose mean is the count predicted by the risk projection model. We will then re-fit the Poisson model to the resulting case counts and use the re-fitted models to predict cancer rates. This approach will provide variance estimates that account for all sources of variation (under some mild assumptions). We will provide a detailed table of the baseline cancer rates that were used for our calculations in the final manuscript. Each factor used in the calculation of the excess number of cancers per 1,000 people has an associated uncertainty, including radiation doses, parameter values of dose-response models, and other factors used for the risk projections. The uncertainty of each component will be described using probability distribution functions and the error propagated by Monte Carlo methods as in previous studies we have conducted (Drozdovitch et al., 2015; Land et al., 2010; Land et al., 2015).

B.3 Methods to Maximize Response Rates and Deal with Nonresponse

Ability to gain and retain participation in study is crucial to gathering accurate data on lifestyle and diet for this group during the 1940's. Advanced community outreach to raise awareness about the project and to encourage participation will be utilized. Invitation letters along with supporting materials from recognized leaders and organizers within the local community will be distributed to potential participants in advance of making contact to request their participation in the study.

To maximize response rate study investigators will also create flyers (**Attachment 4**) that will be distributed to each community, recruiting potential participants. In addition, they will be using community health research networks within each tribal nation and with the Latino community. They will be working with *promotoras de salud* (community health workers) in these networks who will identify potential participants. This proposed recruitment approach has worked successfully on previous studies with these communities. NCI will establish a contract with a university researcher who has extensive experience working with the communities and organizations of interest.

Interviews and focus groups will be conducted at community locations in Albuquerque and New Mexico tribal communities that are frequented by (and convenient to) members of the intended study population. These may include tribal elderly centers, tribal cultural centers, Hispanic/Latino community centers, and/or urban senior centers. Efforts will be made to ensure participation is easy and non-burdensome as possible.

Participation rates and retention will be maximized through several other means such as: providing incentives, extensive moderator training, and consent forms **(Attachment 15).** Discussion and interview materials will be available in Spanish, English, and other native languages as required. The study investigators anticipated a 50% response rate. It is anticipated that 50% of those screened will participate in either the focus groups or an individual interview.

B.4 Test of Procedures or Methods to be Undertaken

In 2014, a successful pilot study was conducted, via individual interviews, collecting information on lifestyle and diet for 9 key informants of Native Americans and Hispanics\Latinos living in New Mexico who were alive at the time of the bomb (currently ≥ 70 years old). More information on the pilot study is found in **Attachment 3**. The pilot study established collaborations and partnerships in the target communities, identified the logistics needed to conduct focus groups, and proved the feasibility of recruiting Hispanic and Native American participants aged 70+, who were alive at the time of the Trinity bomb. Moreover, the data from the pilot phase was used to inform the design of this larger study, which will collect lifestyle and diet information via a series of targeted focus groups. The information acquired from these 9 individual interviews was used to refine the focus group guide (**Attachment 9**).

B.5 Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Individuals who were consulted on the statistical aspects of the study and individuals who will be involved in collecting and/or analyzing data are shown in **Attachment 17.**