

Agricultural Health Study



Demographic, Registry, and Biospecimen Files Users Manual

Release AHSREL201209.00

Stanley E. Legum, Ph.D., Editor

Michael C. R. Alavanja, Dr. P.H., NCI Project Officer

Dale P. Sandler, Ph.D., NIEHS Project Officer

September 2012

TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
1. INTRODUCTION	1
2. DATA COLLECTION PROCEDURES	4
2.1 Buccal Cell Collection.....	4
2.2 Cancer Registry Data.....	5
2.3 Mortality Data	6
2.4 Demographic Data.....	7
2.4.1 Smoking Status Variables	7
2.4.2 Take Home Questionnaire Date.....	11
3. DATA FILE DESCRIPTIONS.....	11
3.1 Buccal Cell Receipt File.....	12
3.2 Cancer Registry File.....	13
3.3 Mortality Data File	14
3.4 Demographic Data File	14
4. CODEBOOKS	15
5. LAB PROCEDURES.....	16
5.1 First Laboratory Protocol	16
5.2 Second Laboratory Protocol.....	19

TABLES

Table 1-1. Composition of Cohort	1
Table 2-1. Resolved Smoking Status Variables.	8
Table 2-2. Smoking Variables.....	9
Table 3-1. Distribution of Applicator Types in Buccal Cell Receipt File.....	12

1. INTRODUCTION

The Agricultural Health Study is a collaborative effort involving the National Cancer Institute (NCI), the National Institute of Environmental Health Sciences (NIEHS), and the U.S. Environmental Protection Agency (EPA). The goals are to investigate the effects of environmental, occupational, dietary, and genetic factors on the health of the agricultural population. This study is providing information that agricultural workers can use in making decisions about their health and the health of their families.

The study has four major components:

1. The main prospective cohort study - cancer and noncancer outcomes:
 - a. Linkage with cancer registries, vital statistics, and the United States Renal Data System (USRDS);
 - b. Ongoing data collection (i.e., telephone interview, food frequency questionnaire, and cheek cell collection (buccal cells));
2. Cross-sectional studies—including questionnaire data, functional measures, biomarkers, and geographic information system (GIS);
3. Nested case-control studies; and
4. Exposure assessment and validation studies.

The cohort includes 89,655 private pesticide applicators, spouses of private applicators, and commercial pesticide applicators recruited within Iowa and North Carolina (Table 1-1). Phase I, initial cohort recruitment, began in December 1993 and concluded in 1997. Phase II follow up began in 1999 and concluded for private applicators and spouses in 2003. Phase II follow up of commercial applicators started in October 2003 and concluded in October 2005. The Phase III followup began in November 2005 and concluded in February 2010. The Phase IV follow up is scheduled to start in 2013.

Table 1-1. Composition of Cohort

Type of Respondent	Number Enrolled
Private Applicators	52,394
Spouses	32,345
Commercial Applicators	4,916
Total	89,655

This study explores potential causes of cancer and other diseases among farmers and their families and among commercial pesticide applicators. Current medical research suggests that, while agricultural workers are generally healthier than the general United States population, they may have higher rates of some cancers, including leukemia, myeloma, non-Hodgkin's lymphoma, and cancers of the lip, stomach, skin, brain, and prostate. Other conditions, such as asthma, neurologic disease, and adverse reproductive outcomes may also be related to agricultural exposures. The Agricultural Health Study is designed to identify occupational, lifestyle, and genetic factors that may affect the rate of diseases in farming populations.

Iowa and North Carolina were selected for this study based on a nationwide competition. Both states have strong agricultural sectors with diverse production methods, commodities, and products. Information we learn from these two states will be helpful to farmers throughout the United States and other countries using modern agricultural technologies.

Phase I data collection involved administration of questionnaires to pesticide applicators and spouses of private pesticide applicators (i.e., spouses of farmers) to obtain information on pesticide use, other agricultural exposures, work practices that modify exposures, and other activities that may affect either exposure or disease risks (e.g., diet, exercise, alcohol consumption, medical conditions, family history of cancer, other occupations, and smoking history).

Phase II data collection involved administration of a computer-assisted telephone interview (CATI) to cohort members approximately 5 years after their initial enrollment. Following completion of the interview, participants were invited to participate in a collection of buccal (cheek) cells and asked to complete a Diet History Questionnaire. Those who agreed to the buccal cell collection were mailed a kit containing mouthwash, a collection cup, an informed consent form, and a preaddressed, postage-paid mailing envelope. Those who agreed to complete the Diet History Questionnaire were sent a questionnaire and a preaddressed, postage-paid mailing envelope.

Phase III data collection also involved administration of a computer-assisted telephone interview (CATI) to cohort members approximately 5 years after completion of a Phase II interview. The focus of the interviews was to update participants' exposure and health information. The sampling frame for this effort excluded members of the cohort who fall into one or more of the following categories:

1. Participant completed the Enrollment Questionnaire only, the spouse of the participant is not enrolled, and nothing else was completed during Phase I or Phase II of the study
2. The participant previously refused all contact from the Agricultural Health Study
3. The participant is deceased.

All participants who do not fit into one of the above categories were eligible for the CATI interview in Phase III. This includes all spouses enrolled in the study. Commercial and private applicators are eligible for the study if they: a) completed more than the Enrollment Questionnaire or had their spouse enrolled in the study; b) did not refuse all contact from the Agricultural Health Study; and c) are alive.

Selected participants were asked to provide buccal cell samples using the same collection methods as in Phase II: Those who agreed to the buccal cell collection were mailed a kit containing mouthwash, a collection cup, an informed consent form, and a preaddressed, postage-paid mailing envelope. Only kits with a single buccal cell collection cup were distributed.

Cohort members were tracked passively through state vital statistics offices and the National Death Index (NDI) to determine vital status and, in the case of deceased persons, cause of death. They were also tracked through the Iowa and North Carolina cancer registries in order to record all malignant cancers in the cohort that were reported to the registries. Addresses for individuals not known to be dead and not contacted in the previous year were requested from the Internal Revenue Service. This was done to verify that they were alive and to determine whether they had moved out of Iowa and North Carolina so that person years of followup could be determined.

This manual describes the data files that are updated annually and are not phase specific. It provides the basic information that an analyst needs to use these files. It includes a brief description of the data collection and editing procedures. Detailed codebooks describing each of the variables and the meanings of the recorded responses are provided as stand-alone documents. The codebooks also contain frequency distributions for each of the variables.

2. DATA COLLECTION PROCEDURES

2.1 Buccal Cell Collection

After a respondent completed the Phase II CATI interview, he or she was asked to provide a sample of buccal cells which could be used for DNA analyses. The interviewer explained that the procedure involved rinsing the mouth with mouthwash, spitting it into a cup, and mailing it back to the study's buccal processing laboratory. If the respondent agreed to participate, he or she was sent a kit containing an informed consent form, an instruction sheet, a small bottle of Scope mouthwash, a cup for collecting the mouthwash after it had been used to rinse the mouth, and a preaddressed postage-paid mailing envelope. From May 2001 through the end of Phase II, respondents who were known to have cancer were sent kits with two cups instead of one. Those receiving these dual buccal kits were given instructions that requested that they use one on each of two consecutive days. A total of 305 kits containing two cups with buccal samples were received between August 2001 and February 2005.

During Phase III, buccal kits were mailed to selected cohort members with cancers diagnosed after enrollment. The cancers used as selection criteria for mailing included:

- Prostate cancers diagnosed from enrollment through 2006 for Iowa cohort members and through 2008 for North Carolina members;
- NHL cases diagnosed after enrollment into 2008;
- Leukemia cases diagnosed starting in January 2006 into 2008; and
- Lung cancers diagnosed starting in January 2006 into 2008.

Only one collection cup was included in each kit mailed during Phase III.

During Phase IV, beginning in July 2012, buccal kits were mailed to cohort members known to have cancers who had not already provided a buccal sample. Priority was given to people with the following cancers: NHL, Bronchus and lung, Colon, Bladder, Leukemia, Rectum.

Respondents placed the completed informed consent form and the buccal collection cup or cups in the mailer and sent them to the processing laboratory in Rockville, Maryland. The laboratory noted the receipt of the consent form and the sample in a receipt file and recorded the amount of material in the specimen. Prior to August 2001, this amount was recorded to the

nearest milliliter. Beginning on August 16, 2001, a flag indicating the volume range was used instead. The flags were:

- H (high – above 17.5 ml)
- N (normal – 10.4 ml to 17.5 ml)
- L (low – 6.0 ml to 10.3 ml)
- V (very low – 1.0 ml to 5.9 ml)
- Z (zero – less than 1.0 ml).

The volume flags were automatically calculated for all those specimens received prior to August 16, 2001 which had exact measurements recorded, so that a volume flag exists for each cup received. A separate volume flag was stored for each of the cups when dual buccal samples were received.

The lab centrifuged the specimen in each cup separately to extract the buccal cells from the mouthwash solution. The pellet of buccal cells was resuspended in Tris EDTA buffer solution and aliquotted into two vials with approximately 1 to 1.5 ml of the cell suspension. Beginning in April 2001, two vials of the supernatant which remained after the buccal cells from each specimen cup had been centrifuged were also stored. Both buccal cells and supernatant were stored at – 70 degrees centigrade.

Detailed laboratory procedures are given in Chapter 5. The Buccal Cell Receipt File includes one record for each buccal kit returned by a respondent whether or not it contained a usable sample. Freezer location data for each vial are available through the National Cancer Institute's Biological Specimen Inventory-II system (CBSI-II).

2.2 Cancer Registry Data

Cancer Registry data are obtained annually from the Iowa and North Carolina cancer registries by the field stations in each state. Each field station submits a list of IDs to its respective cancer registry late each year and then reviews the matches returned to ensure that the data entered into the AHS files represent members of the cohort. Once the files have been adjudicated, they are forwarded to the Coordinating Center and combined into a single data file. The site recode variable (SITEREC) was then added to the dataset using the Site Recode ICD-O-2 (1/27/2003) table and the Site Recode ICD-O-3 (1/27/2003) table provided by NCI's

Surveillance, Epidemiology and End Results (SEER) program. SEER documentation states, “The values of SEER site recode variables are based on the primary site and histology data fields submitted by the registries. The site recode variables define the major cancer sites that are commonly used in the reporting of cancer incidence data.”

In 2010, the state cancer registries also provided Gleason scores for men diagnosed with prostate cancer. These scores were reviewed by staff at NCI in conjunction with the other cancer registry data for these cases. This review led to the creation of derived grade. These values are captured in the variable Grade_GS. The Grade_GS variable reflects changes in the rules for coding prostate cancers diagnosed as of January 1, 2003. It differs from the variable Grade only for prostate cases diagnosed prior to 2003 and for a few post 2002 cases in which the medical record information conflicted with the Grade code. Thus the following two variables have non missing values only for records representing diagnoses of prostate cancer:

- Gleason_Score -- Gleason score for prostate cancer cases
- Grade_GS -- Derived grade for prostate cancer cases

2.3 Mortality Data

Mortality data are obtained annually from the Iowa and North Carolina vital statistics offices by the field stations in each state. Each field station submits a list of IDs to its respective state office late each year and then reviews the matches returned to ensure that the data entered into the AHS files represent members of the cohort. Each field station also submits a list of all cohort members not known to be deceased to the Coordinating Center. These lists are combined and submitted to the National Center for Health Statistics to compare to the latest available annual National Death Index (NDI). The NDI results for each field station are returned to the field station by the Coordinating Center. They are then adjudicated and added to the state’s mortality file. The NDI searches ensure that deaths of cohort members who die out of state but within the United States are reflected in the AHS Mortality File.

Once the state mortality files have been completed, they are forwarded to the Coordinating Center and combined into a single data file.

2.4 Demographic Data

The Demographic data come from a number of sources. They represent the best available demographic and administrative information the project has about a cohort member. Until the end of Phase 3, the administrative cohort data were maintained by Iowa and North Carolina field stations. In June 2011, at the end of the Phase 3, the field stations transferred these data to Westat and Westat set up a SQL server database to store and maintain these cohort data. Some variables in the Demographic File such as date of birth, in state status, and date of move in or out of state are maintained in the cohort database and are updated whenever new information becomes available. Other data (e.g., race and ethnicity) were collected at enrollment through either the Enrollment Questionnaire (for applicators) or the Spouse Questionnaire (for spouses of private applicators).¹ Variables representing which questionnaires have been completed are derived from the latest version of the corresponding AHS data file. Cancer and mortality status variables in the Demographic File reflect data in the Cancer Registry and Mortality Files. If a participant has a record in either the Cancer Registry File or the Mortality File, the corresponding variable in the Demographic file is set to 1 (meaning “Yes”). The last known source and last follow-up date variables are derived using the most recent information available whether from substudies, the outcome of the annual study update mailing,² or the information received from the field stations prior to the end of Phase 3.

2.4.1 Smoking Status Variables

A number of respondents provided inconsistent information about their smoking status in response to smoking status questions asked in the questionnaires in each of the three phases of data collection. In addition, data collected in Phase 3 identified individuals who smoke; but do so on an irregular basis. Data from Phase 1, Phase 2, Phase 3, and guidelines from investigators at NCI and NIEHS for inconsistent and irregular smokers were used to create a consistent set of “resolved” smoking status variables, which have been included on the Demographic File. These resolved smoking variables are listed in Table 2-1.

¹ In a few cases where demographic information was not provided at enrollment, it has been obtained from other sources such as the National Death Index (NDI) or state cancer registries.

²When an envelope was returned because the cohort member no longer lives at that address, we used the new address provided by the post office, when available.

Table 2-1. Resolved Smoking Status Variables.

Variable Name	Description
P1_SMOKE_STATUS_R	Phase 1 resolved smoking status
P2_SMOKE_STATUS_R	Phase 2 resolved smoking status
P3_SMOKE_STATUS_R	Phase 3 resolved smoking status

For purposes of creating the resolved smoking status variables, the respondents were considered in three groups:

- Inconsistent smokers -- Cohort members who gave inconsistent responses to smoking questions across the three data collection phases;
- Irregular smokers -- Cohort members who indicated in the Phase 3 questionnaire that they smoke but indicated that they have never smoked at least 3 cigarettes per week for at least 6 months; and
- Other participants.

Each group is discussed separately below.

2.4.1.1 Inconsistent Smokers

There are 4,887 smokers who have inconsistent smoking responses across the three phases. These responses all involve a response indicating that the respondent never smoked on one of the questionnaires while indicating on another questionnaire that he or she was either a current or former smoker. When the “Never” response came after a “Current smoker” or “Former smoker” response, it is clearly erroneous. In some cases, when the “Never” response came before a “Current smoker” or “Former smoker” response, the age at which a person started smoking is available from the Phase 3 questionnaire. When this is the case, it is possible to determine whether the “Never” response was given before or after the respondent’s reported initial smoking age and to set the resolved smoking status variables accordingly.

In some cases correcting a “Never” response was straight forward. For instance when a respondent indicated that he or she was a former smoker in the Phase 1 and Phase 3 questionnaire responses but responded “Never” during Phase 2, the resolved Phase 2 smoking status variable was set to “Former.” In other cases, it was necessary to review the full set of smoking status and smoking detail responses to determine a resolved smoking status.

The final resolved smoking statuses for these participants were decided using general principles that (1) smoking status cannot be changed from current or former to never and

(2) smoking status for a phase cannot be determined for participants who did not complete a questionnaire for that phase. The variables listed in Table 2-2 were used for setting the resolved smoking status for this class of smokers.

Table 2-2. Smoking Variables.

Variable	Source	
	Data File	Description
SMOK100	Phase 1	During your lifetime, have you smoked at least 100 cigarettes?)
SMOKNOW	Phase 1	Do you smoke cigarettes now?
SMOKPDY	Phase 1	On the average, how much do you or did you smoke each day?
SMOKYRS	Phase 1	What is the total number of years you smoked cigarettes?
A_SMOKE_STATUS	Phase 1	Phase 1 smoking status
TOBACO1	Phase 1	Which of the following tobacco products have you used on a regular basis for six months or longer? - (Pipe)
TOBACO2	Phase 1	Which of the following tobacco products have you used on a regular basis for six months or longer? - (Cigars)
TOBACO3	Phase 1	Which of the following tobacco products have you used on a regular basis for six months or longer? - (Cigarillos)
TOBACO4	Phase 1	Which of the following tobacco products have you used on a regular basis for six months or longer? - (Chewing Tobacco)
TOBACO5	Phase 1	Which of the following tobacco products have you used on a regular basis for six months or longer? - (Snuff)
TOBACO6	Phase 1	Which of the following tobacco products have you used on a regular basis for six months or longer? – (Never used tobacco products)
P2_SMOKE_STATUS	Phase 2	Phase 2 smoking status (Never a smoker; Currently a smoker; or Former smoker)
P3_SMOK100	Phase 3	Have you smoked a total of 100 cigarettes or more during your lifetime?
P3_REGSMOK	Phase 3	Have you ever been a regular smoker, that is have you ever smoked at least 3 cigarettes/week for at least 6 months or more?
P3_SMOKAGESTART	Phase 3	How old were you when you first started smoking cigarettes on a regular basis? By regular basis, I mean 3 cigarettes/week for 6 months or more.

Variable	Source	
	Data File	Description
P3_SMOKYRSTART	Phase 3	What year did you begin smoking?
P3_SMOKFREQ	Phase 3	Do you currently smoke every day, some days or not at all?
P3_SMOKAGESTOP	Phase 3	How old were you when you last smoked cigarettes on a regular basis? By regular basis, I mean 3 cigarettes/week for 6 months or more
P3_SMOKYRSTOP	Phase 3	What year did you last smoke cigarettes on a regular basis?
P3_REGSMOKFREQ	Phase 3	When you used to regularly smoke, did you smoke every day or just on some days?
P3_SMOKNUM	Phase 3	Thinking about all the years that you smoked, about how many cigarettes or packs per day did you usually smoke on days when you smoked? Please tell me whether you are reporting the number of cigarettes per day or the number of packs per day.
P3_SMOKUNIT	Phase 3	Thinking about all the years that you smoked, about how many cigarettes or packs per day did you usually smoke on days when you smoked? Please tell me whether you are reporting the number of cigarettes per day or the number of packs per day. One pack usually equals 20 cigarettes.
P3_PLUS1YRNOSMOK	Phase 3	Thinking about the years between ^StartSmoke and ^StopSmoke, was there ever a period of one year or longer during which you did not smoke cigarettes regularly?
P3_NOSMOKNUM	Phase 3	During the years between ^StartSmoke and ^StopSmoke, for how many total months or years, did you stop smoking cigarettes?
P3_NOSMOKUNIT	Phase 3	During the years between ^StartSmoke and ^StopSmoke, for how many total months or years, did you stop smoking cigarettes?

2.4.1.2 Irregular Smokers

There are 1,663 participants (1,147 private applicators and 516 spouses) who indicated in the Phase 3 interview that they smoke, but not regularly. For these irregular smokers, data from the Phase 1 and Phase 2 interviews were used to establish the Phase 3 resolved smoking status (current/past/never) rather than the Phase 3 questions about current and past

smoking status and cigarette smoking history. Irregular smokers who were current smokers at enrollment and current smokers at Phase 2 were considered to be current smokers at the end of Phase 3. Individuals who were past smokers at enrollment and Phase 2 were considered past smokers at Phase 3. Individuals who were past smokers at Phase 1 and current at Phase 2 were considered current smokers at Phase 3, as these smokers may start and stop smoking multiple times. For the remaining response patterns, the variables displayed in Table 2-2 were used to determine the resolved smoking variable values. For these participants, the final resolved smoking status were decided by using the general constraint that the resolved smoking status cannot be set to Never for current or former smokers.

2.4.1.3 Other Participants

For all other participants the original Phase 1, 2 and 3 smoking status variables were used to populate the corresponding resolved smoking status variables.

2.4.2 Take Home Questionnaire Date

The Take Home Questionnaire date (variable P1TakeHomeQuexDate) is a value entered by the respondent on the Take Home Questionnaire by filling in a matrix of bubbles that were read by an optical scanner. In the case of 81 participants who had entered dates prior to the start of enrollment in the AHS, the Take Home Questionnaire dates were set to the same date as their enrollment dates, which is the date that they completed the Enrollment Questionnaire. The date changes were also made in the Phase I Private Applicator File.

3. DATA FILE DESCRIPTIONS

The following files represent the Demographic, Registry and Biospecimen data:

- Buccal Cell Receipt File
- Cancer Registry Data File
- Mortality Data File
- Demographic Data File

Each of these files can be linked to each other and to the Phase I, Phase II, and Phase III questionnaire data files using the study IDs of the respondents, that is, a combination of PSARTID (the household identifier) and AP_SPOUSE (the indicator of whether the participant is a private applicator, a commercial applicator, or the spouse of a private applicator). A separate codebook exists for each of these files that provides a detailed description of its contents. Each of the files is discussed briefly in a separate section below.

3.1 Buccal Cell Receipt File

The Buccal Cell Receipt File contains the receipt records of buccal cell collection kits sent to the processing laboratory. Kits were entered into the receipt file whether or not they contained usable samples.³ The file provides the date that a collection kit was received, information on whether a completed informed consent form was received for the specimen, a flag representing the gross volume of the submitted sample, an indicator for whether the sample was damaged, a flag indicating whether one or two collection cups were sent by the subject, and comments about the sample or the consent form as needed.

There are 36,358 records in the file representing buccal collection kits that were received between January 22, 1999 and September 19, 2012. This file represents all the kits returned during Phase II and Phase III as well as the kits returned so far during the Phase IV data collection.

The distribution of records in the file is shown in Table 3-1.

Table 3-1. Distribution of Applicator Types in Buccal Cell Receipt File.

Type of Applicator	Number of Buccal Kits Returned
Private applicators	19,924
Spouses of private applicators	14,542
Commercial applicators	1,892
Total	36,358

Because some people who asked to be sent replacement kits eventually sent in both the original kit and the replacement, the total number of unique IDs in the buccal receipt file is 35,333.

³ Kits that were returned by the small number of cohort members who decided after receiving the kits that they did not wish to provide a buccal sample have been excluded from the file.

3.2 Cancer Registry File

The Cancer Registry File contains data on cohort members with cancer. The Iowa and North Carolina field stations searched their state cancer registries to identify records that matched members of the AHS cohort. This information was collected by the Iowa and North Carolina cancer registries and identified cancers diagnosed and the corresponding diagnosis dates.

The State Health Registry of Iowa is a member of the Surveillance, Epidemiology, and End Result (SEER) program sponsored by the National Cancer Institute, which collects data from nine U.S. geographic areas to provide a representative sample of cancer in the United States. For this reason, the Iowa registry conforms to the SEER procedures and standards. Since cancer diagnosis dates are reported to the SEER as month and year only, the diagnosis dates for Iowa have been set to the 15th of the month in the AHS Cancer Registry File.

The North Carolina Central Cancer Registry is a state population-based registry that is not part of the SEER system. This registry receives and reports complete diagnosis dates. An individual may be represented by more than one record because of having multiple cancers. Each record contains basic demographic information about the participant and cancer information including the diagnosis date; a flag indicating whether or not the cancer occurred after enrollment in the study; and ICD codes for cancer site, laterality, morphology, and up to 3 tumor markers. The records also contain information on each cancer's behavior, stage, and grade as well as the SEER site recode variable.

The tumor marker variables are prognostic indicators for specific sites or histologies. These codes are provided for cases diagnosed in 1996 and later. In some cases the markers are more narrowly defined than the type of cancer. For example, for breast cancer, tumor marker 1 is estrogen receptor status; tumor marker 2 is progesterone receptor status. For prostate cases, on the other hand, tumor marker 1 records acid phosphatase (PAP). It records the highest PAP before any prostate treatment. See the SEER Program Code Manual (http://www.seer.cancer.gov/manuals/historic/codeman_1988.pdf) for more details.

There are 12,275 records in the file. The data represent nearly complete cancer diagnoses through 2008 and some 2009 (N = 574) diagnoses.

3.3 Mortality Data File

The Mortality Data File contains dates and causes of deaths of AHS cohort members. The Iowa data were obtained by the Iowa field station's matching records to those of the Iowa Department of Public Health Bureau of Vital Records data files. The North Carolina data were obtained by the North Carolina field station's matching records to the Detailed Death Master file compiled by the North Carolina Center for Health Statistics Vital Records Unit. Both the Iowa and North Carolina data contain death records through calendar year 2009. The file also contains some (N = 25) Iowa death records from 2010.

The Iowa and North Carolina state data were supplemented by reviewing the results of a National Death Index (NDI) search. The NDI search included data from all U.S. states through December 2008. There are 7,866 records in the file.

3.4 Demographic Data File

The Demographic File contains basic demographic information about member of the cohort including birth date, gender, race, ethnicity, highest level of schooling (at enrollment), marital status at enrollment, vital status, in-state status. It also contains a flag indicating whether a person has been diagnosed with cancer and an indicator of whether this is an incident cancer or not. For deceased persons, the file contains the date of death and the date of last contact by the field station.

The file also contains a number of administrative variables including enrollment date, age at enrollment, flags indicating whether or not the individual completed each of the possible Phase I questionnaires, the Phase II and Phase III interview status code assigned by the field station, Phase II and Phase III CATI start date, age on the day that Phase II and Phase III CATI was started, and a flag indicating whether or not a respondent has data in the Supplemental Pesticide file for Phase II or Phase III. Having data in the Supplemental Pesticide file phase indicates that information was collected on the person's use of specific pesticides during that phase of the study.

Some variables such as birth date or gender appear in other data files as well as in the Demographic File. When these differ, the data in the Demographic File are accurate and the other values are what were provided at the time of the interview. These have been preserved in

the CATI interview data files because they are necessary for understanding the branching choices in the CATI interview.

There are 89,655 records in the file.

4. CODEBOOKS

Codebooks for each data file are available as stand-alone documents. For each variable, they list the variable name, indicate whether it is character or numeric, list the question for items drawn directly from the questionnaire or provide a description for derived variables, and give the range of valid values for each variable. For categorical variables, the codebook also specifies the meaning of each value. Frequency distributions are provided for each variable.

The codebooks are:

- Buccal Receipt File
- Cancer Registry File
- Mortality File
- Demographic Data File

5. LAB PROCEDURES

The project used one laboratory to process the buccal specimens from the inception of buccal processing collection through August 15, 2001, and changed to a new laboratory effective August 16, 2001. The protocol used by the first laboratory is included in Section 5.1. That of the second laboratory is included in Section 5.2.

5.1 First Laboratory Protocol

Note: The formatting and numbering of the procedural steps below is that used by the laboratory.

Specimen Processing of Buccal Cells for the Agricultural Health Study

Purpose

This document outlines the procedures for the processing of the buccal cell samples according to the guidelines of the Agricultural Health Study

Definitions

AHS: Agricultural Health Study

MI: milliliter

RLIMS: Relational Laboratory Information Management System

BSC: Class II Biological Safety Cabinet

EDTA- Tris Buffer

Scope

This procedure will be followed by all laboratory personnel handling specimens for the Agricultural Health Study.

Specimen(s)

The specimens that are collected for this study are from farmers who live in two different Geographical locations (Iowa and North Carolina). These farmers are given a kit with a small plastic container and a small bottle of mouthwash. The farmers brush their teeth and then rinse their mouth with the mouthwash and deposit it into the container that is provided. In this specimen, buccal cells are collected

Supplies, Equipment, Reagents, Media - required for testing

- 1) 10 ml pipettes
- 2) 5 ml pipettes
- 3) ml pipettes
- 4) 15 ml centrifuge tubes
- 5) EDTA Tris- Buffer
- 6) 2 ml Cryovials
- 7) Freezer boxes
- 8) Freezer grids
- 9) Labels
- 10) Sample location sheets
- 11) Cryovial racks
- 12) Juoan Centrifuge

Specimen Procedures

Accessing Buccal Samples

- 1) The samples are sent to the laboratory from two sites, Iowa and North Carolina, via US mail.
- 2) Upon arrival, samples are taken into walk-in freezer for storage prior to processing.
- 3) When the sample is taken into the laboratory the envelope is opened to check for the participant's consent forms.
- 4) A barcode is placed on the envelope in which the sample came such that the sample can be scanned into RLIMS.
- 5) While still in the envelope, the sample is then put in a bin so that they can be processed in batch.
- 6) When ready to process, the samples are then taken to the computer and logged into the computer by scanning the bar codes.
- 7) Scan sample into RLIMS
- 8) Save the information that was scanned in.
- 9) Once all the samples have been scanned into RLIMS they are ready to be processed.
- 10) The samples are kept in the walk-in freezer until they are ready to be processed.
- 11) The barcode that is on the envelope can be taken to the computer and labels can be made by using the bar code scanner. 8 labels should be created.
- 12) Once labels are made for all of the samples, you take the samples to the walk-in freezer.

Processing the Buccal samples

- 1) The samples are taken out of the walk-in freezer into the laboratory.
- 2) Once in the laboratory, take the samples to the BSC and take the labels out of the bin.
- 3) Label the 15 ml centrifuge tubes and the sample location form.
- 4) Label cryovials.
- 5) Take the sample out of the envelope.
- 6) Match the number on the envelope with the number on the plastic container that has the buccal sample. Ensure the numbers on the envelope and the plastic container match the number that was assigned to the sample.
- 7) Place two labels on the vials for the supernatant and a red insert cap will be placed in the vial. Place a “supernatant” label on these vials.
- 8) Place two labels on the cryovials used to store buccal cells along with another label which says buccal cells.
- 9) Open the plastic container that contains the buccal sample and aspirate the sample three times using a 10 ml pipet.
- 10) Measure out the amount of fluid that is in the container. Record this measurement on the data sheet.
- 11) Take the sample to the Jovan centrifuge in a test tube rack. Program the centrifuge for program for 15 minutes at 500 rpm. Do not let sample sit for more than 10 mins. In the centrifuge after spin.
- 12) Once the sample has finished its cycle in the centrifuge take them all out and transfer to a test tube rack.
- 13) Put the sample into the BSC and take off the supernatant in each tube and put 1.8 mls of supernatant into each of the 2 cryovials that are labeled for supernatant.
- 14) Once the supernatant is discarded, take a 5 ml pipet and get 3 ml of EDTA-Tris Buffer.
- 15) Mix the cells up by aspirating about 4 times.
- 16) Open up the cryovial and dispense 1 ml per cryovial into 2 vials.
- 17) Fill the data sheet out for the cell location.
- 18) Turn data sheet into the data information group.

5.2 Second Laboratory Protocol

Note: The formatting and numbering of the procedural steps below is that used by the laboratory.

Processing of AGRICULTURAL HEALTH STUDY Buccal Cell Samples

- 1.0 Purpose: To describe the receipt, processing and storage of buccal cell samples for the AGRICULTURAL HEALTH STUDY.
- 2.0 Policy: The receipt, processing, and storage will comply with Good Laboratories Practices (GLPs)
- 3.0 Equipment:
 - 3.1 IEC Centra-8 or Forma Scientific Centrifuge
 - 3.2 Biological Safety Cabinet
 - 3.3 Refrigerator
 - 3.4 -80°C Freezer
- 4.0 Materials:
 - 4.1 50 ml conical centrifuge tube
 - 4.2 5.0 ml disposable sterile serological pipettes
 - 4.3 Tris EDTA buffer solution pH 8.0 (Quality Biological Inc. #351-0110131)
 - 4.4 2.0 ml sterile cryovials
- 5.0 Procedure:
 - 5.1 Specimen Receipt Procedure
 - 5.1.1 Open and remove two specimen collection cups from participant's envelope. An Informed Consent form should be included with the specimens. Review and ensure participant has signed and dated form.
 - 5.1.2 Attach a specimen identification label (AG+5) on the Informed Consent Form.
 - 5.1.2.1 The Informed Consent Forms are copied, sorted by site, and sent to sites each week. Receipt of the Informed Consent Forms will be confirmed by the sites and copies will be sent to BioReliance for destruction.
 - 5.1.2.2 An Informed Consent Form may be received separately from the collection cups. In the receiving database, change the Informed Consent Form field to a "Y" to reflect receipt of this form. Sites will notify McKesson by E-mail if they receive an Informed Consent Form. Change the Informed Consent Form field in the receiving database to a "Y" to reflect receipt of this form. E-mail

the site the AG number associated with the Informed Consent Form.

5.1.3 Buccal cell samples for the AGRICULTURAL HEALTH STUDY should be processed on the day of receipt. If this is not possible, the samples are to be stored in the refrigerator until they are processed.

5.1.4 Receive samples by accessing the Biological Inventory System (BioSIS) and selecting the Agricultural Health Study Receiving Module.

5.1.5 Enter the following information into the Agricultural Health Study database :

5.1.5.1 Enter Study Number (AHS Buccal)

5.1.5.2 Enter date envelope received

5.1.5.3 Did participant refuse to provide a specimen (Y or N)?

5.1.5.4 Were dual cups received (Y or N)?

5.1.5.5 Enter the estimated volumes in the collection cups.

Z- None: Below 1.0 ml

V- Very Low: 1.0 ml to 5.9 ml

L- Low: 6.0 ml to 10.3 ml

N- Normal: 10.4 ml to 17.5 ml

H- High: Above 17.5 ml

5.1.5.6 Was the sample damaged or spilled (Y/N)?

5.1.5.7 Was the Informed Consent Form received (Y/N)?

5.1.5.8 Did the participant sign the Informed Consent Form (Y/N)?

5.1.5.9 Enter the date the Inform Consent Form was signed.

5.1.5.10 Enter comments concerning the Informed Consent Form or specimen.

5.1.5.11 Scan the patient barcode on the Informed Consent Form.

5.1.5.12 Scan the barcode(s) on the collection cup(s).

5.1.5.13 Scan the barcode on the Specimen Identification label.

5.1 Specimen Labeling Procedure

5.2.1 Generate labels using the AGRICULTURAL HEALTH STUDY Label Generation Program. Ten labels are generated per participant; five per collection cup received. Label information is listed below:

Label Affix To	Label Information
-----------------------	--------------------------

No #	Centrifuge Tube AG +5 digits and collection cup ID
------	--

# 01	Cryovial- Cells AG +5 digits-1 and collection cup ID
------	--

# 02	Cryovial- Cells AG +5 digits-2 and collection cup ID # 03 Cryovial-Supernatant AG +5 digits-3 and collection cup ID
------	--

# 04	Cryovial-Supernatant AG +5 digits-4 and collection cup ID
------	---

- No # Centrifuge Tube AG +5 digits and collection cup ID
- # 05 Cryovial-Supernatant AG +5 digits-5 and collection cup ID
- # 06 Cryovial-Supernatant AG +5 digits-6 and collection cup ID
- # 07 Cryovial-Cell AG +5 digits-7 and collection cup ID
- # 08 Cryovial-Cell AG +5 digits-8 and collection cup ID

5.2.2 Label the centrifuge tubes and vials.

5.2.3 Move the collection cups and centrifuge tubes into the Biological Safety Cabinet and prepare for processing.

5.3 Specimen Processing Procedure

5.3.1 In the Biological Safety Cabinet, invert the collection cups 4 to 6 times before removing the top. Check labels on the centrifuge tubes. The ID number on each centrifuge tube must match the ID number on the collection cup. Transfer each specimen into a separate centrifuge tube. If a participant sends in two cups, each cup must be placed in a separate centrifuge tube. Continue until all specimens have been transferred.

5.3.2 Place the centrifuge tubes into the IEC Centra-8 or Forma Scientific centrifuge. Centrifuge at 1500 x g for 15 minutes (3100 RPM), with the break off.

5.3.3 Move the centrifuge tubes and the corresponding cryovials into the Biological Safety Cabinet. Before starting, make sure the AG number on the centrifuge tube corresponds to the AG number on the cryovials.

5.3.4 Using a sterile 5 ml pipette, remove 3.6 ml of the supernatant. Place 1.8 ml of supernatant into each of the two cryovials #3 and #4. If applicable, the supernatant from the second collection cup will be placed into cryovials #7 and #8. Change pipettes if a second collection cup is used.

5.3.5 Discard the remaining supernatant from the centrifuge tube.

5.3.6 Using a 5 ml sterile pipette, add 3 ml of Tris EDTA pH 8.0 to the centrifuge tube. Disperse the buccal cell pellet by gently tapping the 50 ml tube and then resuspend using the pipette.

5.3.7 Remove the entire sample from the tube and aliquot equal amounts (approximately 1.5-1.8 ml) into two 2.0 ml sterile cryovials label #1 and #2. If applicable, the cells from the second collection cup will be placed into cryovials #5 and #6.

5.4 Specimen Inventory Procedure

5.4.1 Access the Biological Specimen Inventory System (BioSIS) and select the AGRICULTURAL HEALTH STUDY Inventory Module.

5.4.2 Enter the date of processing into the inventory system.

5.4.3 Scan the barcodes on the cryovials into the inventory system. As the vials are scanned, the computer will automatically assign the location codes.

5.4.4 Place the cryovials into the -80°C freezer.

5.5 Specimen Data File Transfer Process

5.5.1 At the end of each day, generate the receiving file. This file is electronically sent to Westat via FTP. At the end of the month, the entire receiving and inventory databases are sent to Westat via FTP. Send Westat an E-mail message confirming that the files have been transferred.