

Agricultural Health Study



Phase II

Data File Users Manual

Release P2REL0712.02

Stanley E. Legum, Ph.D., Editor

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

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

Phase II Data Working Group

Michael C. R. Alavanja, Dr. P.H.

Joe Coble, Sc.D.

Jane Hoppin, Sc.D.

Stanley E. Legum, Ph.D.

Stuart Long

Dale P. Sandler, Ph.D.

Claudine Samanic, M.S.P.H.

Marsha Shepherd

Cassandra Swartz, M.S., M.B.A.

February 2008

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page</u>
1.	INTRODUCTION	1-1
2.	DATA COLLECTION PROCEDURES	2-1
2.1	Eligibility	2-1
2.2	Main Data Collection.....	2-1
2.2.1	CATI Interviews	2-2
2.2.1.1	Pesticide Use Sections	2-5
2.2.1.2	Health Modules	2-7
2.2.1.3	Cooking Practices Module	2-8
2.2.2	Diet History Questionnaire	2-8
2.3	Supplementary Data Collection	2-9
3.	DATA PROCESSING PROCEDURES	3-1
3.1	Data Cleaning	3-1
3.2	Preparation of Analytic Files	3-2
3.2.1	Variable Naming Conventions.....	3-3
3.2.2	Derived Variables	3-3
3.2.2.1	Conversion of Arrays to Indicator Variables	3-4
3.2.2.2	Variables in Supplemental Files Summarized in Analysis Files.....	3-5
3.2.2.3	Functional and Chemical Class Flags	3-6
3.2.2.4	Conversion of Multiple Response Lists to Indicator Variables	3-6
3.2.2.5	Added Names of Active Ingredients Corresponding to the PC Codes.....	3-15
3.2.2.6	Created Days per Year and Hours per Year Variables.....	3-16
3.2.3	Missing Value Codes	3-16
3.2.4	Treatment of Differences among Versions of Modules.....	3-16
3.2.4.1	Differences between Versions 1 and 2 of Private Applicator Pesticide Use Module	3-17
3.2.4.2	Differences between Versions 2 and 3 of the Private Applicator Pesticide Use Module	3-18
3.2.4.3	Differences among Versions of the Male Health Module.....	3-22
3.2.4.4	Differences among Versions of the Female Health Modules	3-22
3.2.4.5	Differences among the Versions of the Cooking Practices Module	3-22
3.3	Pesticide Exposure Flags	3-23
3.4	Pesticide Exposure Algorithms.....	3-25
3.4.1	Average Intensity Score.....	3-28
3.4.2	Mix, Apply, and Repair Scores.....	3-28
3.4.2.1	Mix Score.....	3-29
3.4.2.2	Apply Score	3-29
3.4.2.3	PPE Factors	3-36
3.4.2.4	Repair Score	3-38
3.5	Version 2 General Considerations	3-38
3.6	Body Mass Index Estimates.....	3-45
4.	DATA FILE DESCRIPTIONS.....	4-1
4.1	Private Applicator File.....	4-1
4.2	Private Applicator Supplemental Pesticide File.....	4-6

4.3	Spouse File.....	4-7
4.4	Spouse Supplemental Pesticide File	4-8
4.5	Commercial Applicator File	4-8
4.6	Commercial Applicator Supplemental Pesticide File	4-8
4.7	Diet History Questionnaire File	4-8
5.	USAGE NOTES	5-8
5.1	How to Link Files for Analyses	5-8
5.2	Discrepancies between Phase I and Phase II Files.....	5-8
5.3	Unusual Values	5-8
5.4	Interpretation of Missing Data Patterns	5-8
5.5	Topics Discussed in Phase I User Manual that May Be of Interest	5-8
6.	REFERENCES	6-8
7.	QUESTIONNAIRES	7-8
8.	CODEBOOKS	8-8

List of Appendixes

Appendix

A	Pesticides Included on A List	A-1
B	Pesticides Included on B List.....	B-1
C	BMI Estimation Project	C-1

TABLE OF CONTENTS

List of Tables

<u>Table</u>	<u>Page</u>
Table 1–1. Composition of Cohort.....	1-1
Table 2–1. Number of private applicators completing each version of the Pesticide module by state.....	2-6
Table 2–2. Number of spouses completing each version of the Pesticide module by state.....	2-6
Table 3–1. Functional classes	3-8
Table 3–2. Chemical classes	3-9
Table 3–3. Indicator variable names associated with each Response Option to Question 20a in the Phase II Private Applicator Questionnaire.....	3-10
Table 3–4. Indicator variable names associated with each Response Option to Question 35 in the Phase II Private Applicator Questionnaire	3-11
Table 3–5. Indicator variable names associated with each Response Option to Question 36 in the Phase II Private Applicator Questionnaire	3-13
Table 3–6. Indicator variable names associated with each Response Option to Question 39d in the Phase II Private Applicator Questionnaire.....	3-14
Table 3–7. Indicator variable names associated with each Response Option to Diet Module Question 1a in the Phase II Private Applicator file.....	3-15
Table 3–8. Values of PC code and active ingredient name variables for Pursuit Plus	3-15
Table 3–9. Change in values of methyl bromide indicator variable because of changes in definition of indicators.....	3-25
Table 3–10. Lifetime Indicator Variable through Phase II Logic	3-25
Table 3–11. Lifetime Exposure Days through Phase II Calculation.....	3-27
Table 3–12. Imputed Application Method Names and Scores for Private Applicators with Missing Application Methods.....	3-34
Table 3–13. Imputed Application Method Names and Scores for Spouses of Private Applicators with Missing Application Methods.....	3-35
Table 3–14. Imputed Application Method Names and Scores for Commercial Applicators with Missing Application Methods.....	3-36
Table 3–15. Assignment of Method Codes to Other Specify Responses for Each Question.....	3-39
Table 3–16. Application Method Scores Associated with Each Application Method Code.....	3-40
Table 3–17. Phase II PPE Variables:	3-42
Table 3–18. New Application Method Codes and Scores	3-43
Table 3–19. Recode of Phase II Application Method using P3 Codes	3-44
Table 4–1. Description of administrative variables	4-1
Table 4–2. Indicator variables for functional and chemical class groups	4-3
Table 4–3. Phase I demographic variables included in Phase II Private Applicator file	4-6
Table 4–4. Phase I demographic variables included in Phase II Spouse File	4-7
Table 5–1. Gender and Applicator/Spouse Variable Names and Availability in AHS Data Files	5-8

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 89,658 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 follow up began in November 2005.

Table 1–1. Composition of Cohort

Type of Respondent	Number Enrolled
Private Applicators	52,395
Spouses	32,347
Commercial Applicators	4,916
Total	89,658

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.

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 follow up could be determined.

This manual describes the Phase II data files and provides the basic information that an analyst needs to use them. It includes a brief description of the data collection and editing procedures, usage notes to guide the analyst, copies of the questionnaires completed by the participants, and detailed codebooks describing each of the variables and the meanings of the recorded responses. The codebooks also contain frequency distributions for each of the variables.

2. DATA COLLECTION PROCEDURES

2.1 Eligibility

Farmers, spouses of farmers, and commercial applicators who had enrolled in Phase I of the AHS by completing one or more of the Phase I questionnaires were eligible for the Phase II interviews whether or not they still resided on a farm and whether or not they still participated in agricultural activities. Individuals who completed all or part of the Phase II CATI interview were eligible to participate in the buccal cell and Diet History Questionnaire data collections. Participation in each of the data collection activities was voluntary.

A few cohort members were excluded from the Phase II interviewing and other active data collection activities. These included those who died prior to being contacted for an interview, those too ill to participate, those who requested that they no longer be contacted by the study, those who participated in pilot administrations of the CATI instrument, and those who could not be located. The last group included people for whom no phone number could be found, as well as those with unpublished numbers or blocked lines.

Some individuals died after they were interviewed. Therefore, a researcher should not assume that someone whose ID appears in the Mortality File was excluded from interviewing. The cancer status and vital status of people excluded from the interviews were still tracked passively using the Iowa, North Carolina, and national databases available to the study.

2.2 Main Data Collection

The primary data collection activity during Phase II of the AHS was the administration of CATI interviews. Separate, though related, interviews were planned for each of the three major subgroups within the cohort: private pesticide applicators (henceforth referred to either as private applicators or farmers), spouses of private applicators, and commercial pesticide applicators (henceforth, commercial applicators). Private applicators were interviewed between January 1999 and September 2003. Spouses of private applicators were interviewed between July 1999 and August 2003. Commercial applicators were interviewed between October 2003 and were and October 2005.

2.2.1 CATI Interviews

The Phase II interviews consisted of three basic questionnaires: one for private applicators one for spouses of private applicators, and one for commercial applicators. The questionnaires were similar, with differences attributable to the assumption that applicators and spouses were differently exposed to pesticides and taking into account that spouses were mostly female while applicators were mostly male. The questionnaires went through multiple revisions.

While the private applicator and spouse questionnaires first asked people to list the types of crops and animals raised on their farms and then asked what pesticides were used for each and for noncrop applications (e.g. weed control on fence rows), the commercial applicator interview focused on ways in which the applicator used pesticides. For example, question 14 asked whether the commercial applicator personally mix, load, handle or apply pesticides for:

1. Farm crops
2. Lawns or turf, including cemeteries and golf courses
3. Farm animals/shelters
4. Stored grain or other agricultural products
5. Interior or exterior building applications (home, commercial, or public)
6. Nursery or garden (including shrubs, flowers and ornamentals)
7. Weed or brush control (for fence rows, pastures, public lots, or highway right of ways)
8. Trees, including fruit trees
9. Rodent or termite control
10. Other applications

Questionnaires for the CATI Interview

The CATI interviews, which were written in Blaise, were constructed in a modular fashion. The questionnaires all used the same modular structure and administered the modules in the order: Start, Health, Crops, Animals, NonCrops, Personal Protective Equipment (PPE), High Exposure, and Diet and Cooking Practices.

The first section verified the identity and demographic information of the respondent and asked about the source of drinking water, with some follow up questions if the source was a private well. The section ended by asking if the respondent had been actively engaged in farming in any year since the

year of enrollment and, if so, what was the most recent year. The most recent year in which the respondent farmed was referred to as the “reference year” for that person. Most questions asked about the person’s experience during his or her reference year. At the end of the starting section, the respondent was asked if he or she had personally mixed, loaded, or applied pesticides and, if so, the number of days during the reference year and the average number of hours during a day that he or she used any pesticides.

At this point in the interview there was a major branching point (sometimes referred to as the decision box) that determined which of the paths through the questionnaire a respondent would take. The branching used information on whether the respondent had been farming since enrollment; if so, whether or not the respondent provided the interviewer with a reference year; and whether the respondent personally handled pesticides (i.e., personally mixed, loaded, or applied pesticides).

For the private applicators, the following four major branches of the interview were determined by the information collected during the starting module:

1. If the respondent had been farming, had provided a reference year, and had personally mixed, loaded, or applied pesticides, then all of the modules and all of the questions in the modules were administered (except for some minor skips).
2. If the respondent had been farming and had provided a reference year, but had NOT personally mixed, load, or applied pesticides, then all of the modules were administered but most of the detailed questions about pesticides were skipped.
3. If the respondent answered “No,” “Don’t know,” or “Refused” to the “Have you been farming?” question, OR if the respondent did not provide a reference year (even if he or she had been farming), AND the respondent had personally mixed, loaded, or applied pesticides, the Crop and Animal modules were skipped and the NonCrop module and all following modules were administered.
4. If the respondent answered “No,” “Don’t know,” or “Refused” to the “Have you been farming?” question, OR if the respondent did not provide a reference year (even if he or she had been farming), AND the respondent had NOT personally mixed, loaded, or applied pesticides, all of the modules were skipped except the Health and the Diet and Cooking Practices modules.

For the spouses of private applicators, the decision box generated eight major paths through the questionnaire. This occurred because Question 11 in the Private Applicator Pesticide Use module, which read, “Since (date of enrollment), did you personally mix, load or apply pesticides for crops, animals, or for any other purposes?” was expanded to three questions in the Spouse version of the module:

Spouse Q11a. “Since (date of enrollment), did you personally mix, load or apply pesticides for use on crops? Do not include garden produce unless sold for profit.”

Spouse Q11b. “Since (date of enrollment), did you personally mix, load or apply pesticides for animals or animal confinement areas?”

Spouse Q11c. “Since (date of enrollment), did you personally mix, load or apply pesticides for any other purpose? This would include applications for garden, lawn, nursery, fence rows and buildings. It does not include applications made to pets or in your home.”

The eight paths through the Spouse Pesticide Use Module were as follows:

Path 1: Go through all modules (respondent answered “Yes” to 11a, 11b, and 11c)

Path 2: Go through all modules except the NonCrop module (respondent answered “Yes” to 11a and 11b, but “No,” “Don’t Know,” or “Refused” to 11c)

Path 3: Go through all modules except the Animal module (respondent answered “Yes” to 11a and 11c, but “No,” “Don’t Know,” or “Refused” to 11b)

Path 4: Go through all modules except the Animal and NonCrop modules (respondent answered “Yes” to 11a, but “No,” “Don’t Know,” or “Refused” to 11b and 11c)

Path 5: Go through all modules except the Crop module (respondent answered “Yes” to 11b and 11c, but “No,” “Don’t Know,” or “Refused” to 11a)

Path 6: Go through all modules except the Crop and NonCrop modules (respondent answered “Yes” to 11b, but “No,” “Don’t Know,” or “Refused” to 11a and 11c)

Path 7: Go through all modules except the Crop and Animal modules (respondent answered “Yes” to 11c, but “No,” “Don’t Know,” or “Refused” to 11a and 11b)

Path 8: Go through all modules except the Crop, Animal, and NonCrop modules (respondent answered “No,” “Don’t Know,” or “Refused” to Questions 11a, 11b and 11c)

While the questionnaires consisted of several modules, the survey can usefully be thought of as having just two main sections: Pesticide Use and Health. The content of the pesticide use section varied depending on whether the respondent was an applicator or a spouse. The content of the Health modules varied depending on whether the respondent was male or female and, for women, whether they were pre- or postmenopausal as reported during Phase I of the study. Each of the major sections is discussed separately below.

2.2.1.1 Pesticide Use Sections

Both the Private Applicator and the Spouse Pesticide Use sections had the following modules: Start, Crops, Animals, Noncrops, Personal Protective Equipment, High Exposure, and Diet and Cooking practices. The pesticide use sections collected information on pesticides handled by applicators during the reference year, duration and frequency of use, application methods, personal protective equipment used, personal hygiene (for example, when and where work clothes and boots were changed, how soon hands were washed and showers taken after handling pesticides, how work clothes were washed), high pesticide exposure events, and risk-taking attitudes. The questionnaire collected data on pesticide use for crop, animal, and noncrop (e.g., fence line weed control) applications.

For applicators, the reference year was defined as the last year in which they farmed. For applicators who were still farming at the time of the interview, the interview began in the months of January through October, the reference year was taken as the previous calendar year. If the interview was begun in November or December, the reference year was taken as the current calendar year. For spouses, the reference year was defined as the calendar year prior to the one in which the interview took place. Thus, an applicator and spouse interviewed during the same month would have the same reference year unless the interview was conducted during the months of November or December.

While the Pesticide Use module for spouses was similar to that for private applicators, it focused more on household, lawn, and pet issues. For instance, in addition to questions about farming activities, it asked whether family members who were working in farm areas usually took their boots off before entering the house. It also asked about treatments of the house for termites and other insects and about flea and tick control measures taken to protect pets.

There were three versions of the Private Applicator questionnaire and two versions of the Spouse questionnaire. One of the analysis file creation tasks was to combine the versions within a questionnaire; this process reduced three Private Applicator files to one and two Spouse files to one.

There were three versions of the private applicator Pesticide Use module. They were generally identical and any differences were small.¹ For instance, there were minor changes in wording to clarify the intent of a question, the number of items that could be reported in some “Other, specify”

¹ Note, however, that Version 1 and Version 2 of the Private Applicator Pesticide Use module branched to the male health module, but did not branch to the female health modules.

questions was expanded, and there were minor revisions to routing rules. In a few cases the set of responses to a question was shortened by combining two choices that had essentially the same meaning (e.g., “stored grain” and “grain bin applications”) or adding a new response choice so that fewer responses needed to be collected in an “Other, specify” follow-up question.

The differences in specific questions are discussed in detail in Section 3.2.4. These differences should, for the most part, be transparent to the analyst. In preparing the analytic files, the final version of each questionnaire was taken as the target to be reflected in the data set.

Table 2-1 displays the number of private applicators by state completing each version of the CATI questionnaire. Note that Version 1 of the Private Applicator Pesticide Use module was used briefly in Iowa and was never used in North Carolina.

Table 2–1. Number of private applicators completing each version of the Pesticide module by state

Version	State		Total
	Iowa	North Carolina	
1	79	0	79
2	2,521	1,524	4,045
3	18,489	10,844	29,333
Total	21,089	12,368	33,457

There were two versions of the Spouse Pesticide Use module. The differences between the modules were of the same types and magnitude as those in the Private Applicator Pesticide Use modules. Female health questions were, however, called from both versions of the Spouse Pesticide Use module.

Table 2-2 displays the number of spouses by state completing each version of the CATI questionnaire.

Table 2–2. Number of spouses completing each version of the Pesticide module by state

Version	State		Total
	Iowa	North Carolina	
1	5,329	1,312	6,641
2	11,602	5,553	17,155
Total	16,931	6,865	23,796

Within the Pesticide Use section, the Crop module asked about what crops were grown (up to 10 crops), what pesticides were used for each crop (up to 15 pesticides per crop) and, for each pesticide, how many days and hours were spent mixing, loading, or applying it. Detailed questions about application method and the form of the pesticide were then asked about the “top four” pesticides per crop, as determined by the number of days spent mixing, loading, or applying. The CATI code actually identified the top four values (having no way to choose one pesticide over another if there were ties), and the detailed questions were asked about any pesticide having any of those values. In some cases, there were more than four pesticides per crop.

The Animal module asked what animals were raised (up to six) and which type of animal got the most pesticide applications. Then, for that type of animal only, the module asked what pesticides were used (up to six) and, for each pesticide, the application method used and the number of days and hours spent applying pesticides.

The NonCrop module asked what types of noncrop pesticide applications were done (up to eight or nine, depending on the version), what pesticides were used (up to four pesticides per noncrop application) and, for each pesticide, what application method and how many days and hours were spent applying pesticides.

The Personal Protective Equipment (PPE) module asked if the respondent wore protective clothing (e.g., gloves, Tyvek overalls, a face shield) or used protective equipment such as a tractor with an enclosed cab and a charcoal filter.

The High Exposure module asked if there had been any pesticide incidents giving the respondent a high level of exposure and, if so, what the pesticide and the symptoms were.

2.2.1.2 Health Modules

The Health modules collected data on height, weight, smoking status, alcohol use, and numerous medical conditions. The list of medical conditions varied somewhat for men and women. The Women’s Health modules also varied depending on whether they were pre- or postmenopausal or did not indicate menopausal status when they completed the Phase I questionnaire. Women who did not state

their menopausal status during Phase I were asked additional questions depending on whether they were applicators or spouses of applicators. The five health modules were:

Five Year Follow-up Health Module for Men

Follow-up Health Questionnaire for Women Version A – for women who had not gone through menopause as indicated by an answer of “No” or “Not sure” to Question 6 of the Phase I Female and Family Health Questionnaire.

Follow-up Health Questionnaire for Women Version B – for women who went through menopause before enrollment, as indicated by an answer of “Yes” to Question 6 of the Phase I Female and Family Health Questionnaire.

Follow-up Health Questionnaire for Women Version C: Spouse – for women who completed a Phase I Spouse Questionnaire but did not complete a Phase I Female and Family Health Questionnaire;

Follow-up Health Questionnaire for Women Version C: Applicator – for female pesticide applicators who completed a Phase I Enrollment Questionnaire but did not complete a Phase I Female and Family Health Questionnaire. This questionnaire was also used for spouses prior to April 1, 2000.

Each of these modules is included in the appendices. Only one of the Health modules was completed by each respondent.

2.2.1.3 Cooking Practices Module

The Cooking Practices module was administered at the end of the CATI interview. It collected information on eating and cooking habits for meat, poultry, and fish. A copy of the module can be found in the appendices.

2.2.2 Diet History Questionnaire

After a respondent completed the Private Applicator CATI Questionnaire or the spouse CATI Questionnaire, he or she was also asked to complete a Diet History Questionnaire (DHQ). If the respondent agreed, a questionnaire was mailed out with a pre-addressed postage-paid return envelope. A copy of the DHQ is included in the appendices.

Returned questionnaires were optically scanned and checked for scanning errors. Data from these questionnaires can be found in the Diet History Questionnaire File.

2.3 Supplementary Data Collection

Because of a programming error in Versions 1 and 2 of the spouse CATI instrument (specifically, the Female Health Module, Version B), 311 women who indicated on the Phase I Female and Family Health Questionnaire that they had undergone natural menopause at age 45 or earlier (N=239) or who left their ages at menopause blank (N=72) were not asked a series of follow-up questions in Version B of the Followup Health Questionnaire for Women. This affected questions 22 through 25a in the questionnaire. The interviewing program was changed in September 2000 to correct the problem.

Between November 20, 2001 and January 10, 2002, an attempt was made to call 307 women who should have been asked these questions, to verify the age at menopause and to ask the questions which had been skipped.² Supplementary interviews were completed with 283 (92%) of these women. The responses from this data collection were recorded on paper forms, then key entered and verified. The data from this interviewing effort were integrated into the Spouse data file. A flag (FEMBFLAG) was set to indicate which women completed these questions during the supplemental phone call rather than during the regular CATI interview. These data then underwent the same set of data cleaning operations as the data collected in the normal manner.³

² Four women were excluded from these calls: two had requested no further contact, one could not be interviewed because of a language or hearing problem, and one had died.

³ Five women had listed their age at menopause as less than 17 on the Phase I questionnaire. All of these women indicated during the supplementary interview that their ages at menopause were over 45. Of the 72 women who had left their age at menopause blank on the Phase I questionnaire, 57 were over 45, 11 were under 45, and 4 did not remember their age at menopause. Thus 127 (45%) of the 283 women who completed the supplementary interview were actually over the age of 45 and should not have been asked the questions that were skipped.

3. DATA PROCESSING PROCEDURES

A major advantage of collecting data through computer assisted interviewing is that basic checks can be made on the data as it is entered, giving the interviewer fewer opportunities to make errors (impossible values are blocked) and to correct some types of errors immediately (“soft” range checks warn the interviewer of unusually high or low values). Another important advantage is that data files are available in electronic form as soon as the interview is complete.

The questionnaire contained a list of pesticides from which interviewers could select (see Questions 14, 24a, and 28c in the Private Applicator Pesticide Use module). If the respondent named a pesticide not on the list, it was keyed in by the interviewer as the response to Question 14a, Question 24a1, or Question 28c1 in the Private Applicator Pesticide Use module. The Spouse Pesticide Use module had a similar structure. The field stations periodically reviewed the pesticide names key entered by the interviewers. In some cases, they corrected the spellings entered by the interviewer and in some cases they added the pesticide to the list presented in the pick list. When a corrected spelling or the addition of the pesticide to the pick list meant that an item in the “Other” response (e.g., Q14a) was now in the list, the field stations moved the response to the variable based on the pick list (e.g., Q14).

After completing the editing, the field stations forwarded copies of the interviews to the Coordinating Center. These data files excluded personal identifiers but were otherwise complete. The Coordinating Center converted the files to SAS data files and checked them for errors. Once the files were certified as clean, analytic data files were produced.

The data cleaning activities and analysis file preparation activities are described in separate sections below.

3.1 Data Cleaning

The data cleaning task was performed jointly by the field stations and the Coordinating Center. The field stations performed post-interview data cleaning and editing as described above. After receiving a data file from a field station, the Coordinating Center ran a set of edit checks on the data outside of Blaise and produced a report of any problems for the field station.

The checks replicated all of the range, skip, and logic checks built into the CATI instrument and incorporated a number of additional, more complex checks as well. The Coordinating Center edit checks included:

- **Range checks** – Values that were outside of the valid range for a variable were reported as errors.
- **Forward and backward skip pattern checks** – These verified that skip patterns had been followed, questions that should have been skipped had missing values, and questions were answered only if the associated lead-in questions had response values that allowed the interviewer to ask those questions.
- **Logic checks** – Values among related questions were checked for plausibility.
- **Outlier checks** – Extremely low and high values (determined using Tukey’s (1977) method) were flagged for consideration by the field stations for selected variables that were expected to be used in many analyses.

The field station staff reviewed the edit reports and either made corrections or recorded a decision to allow a data item to stand by setting a reason flag in an override file which was stored in the form of an MS Access database. Cumulative files containing corrections and the corresponding override file were returned to the Coordinating Center. The edit checks were run on each newly arrived file, but items were not included on the report if they had already been flagged as overrides.

3.2 Preparation of Analytic Files

The data from the corresponding questionnaire (i.e., Spouse with Spouse and Private Applicator with Private Applicator) from Iowa and North Carolina field stations were combined. Variables were recoded so that a given value for a variable had the same meaning regardless of which version of the questionnaire contained it. Variable names were assigned so that a given variable name represented the same question content independent of the questionnaire module administered. The data from the different questionnaire versions of the Applicator Questionnaire modules were then combined into a single file, as were the data from the different versions of the Spouse Questionnaire modules.

Values were recoded to make them easier for analysts to use and to make them consistent with the coding used in the Phase I data files. For example, a response of “Yes” was coded as 1 and a response of “No” was coded as 0.

3.2.1 Variable Naming Conventions

Variables from modules administered only to private applicators were prefixed with A2. Variables from modules administered only to commercial applicators were prefixed with C2. Variables from modules administered only to spouses were prefixed with S2. Since many of the questions were repeated in all the Health modules, and since the Male Health module and Version C (applicator) of the Women's Health module were for a period administered to both applicators and spouses, the questions that were common to multiple Health modules were given the same variable name (except for the A2, C2, and S2 prefixes). Since only one Health module was administered during an interview, this naming convention will not cause any ambiguity as to what questions a respondent was asked.

Use of the A2, C2, and S2 prefixes ensured that the Phase II variable names were distinct from the Phase I variable names. The exceptions to this are the variables PARTID (participant ID) and SPSPAIR (spouse pair ID), which have the same name and meaning in all AHS data files in which they appear. In some instances (e.g., the Phase I exposure flags and estimates), variables were copied directly from Phase I files. When this was done, the variable names and labels are identical to their counterparts in the Phase I files.

3.2.2 Derived Variables

Derived variables are variables created by combining or manipulating responses to items asked directly in the questionnaires. A number of types of derived variables are included in the data files:

- Conversion of arrays to indicator variables;
- Variables in supplemental files summarized in analysis files;
- Functional and chemical class flags;
- Conversion of multiple response lists to indicator variables;
- Added names of active ingredients corresponding to the PC Codes; and
- Created days per year and hours per year variables.

The derivation of each type is discussed in a separate section below.

3.2.2.1 Conversion of Arrays to Indicator Variables

The data structures used for data collection during the interviews were effective for that purpose but are awkward to use for the purposes of data analysis. For instance, applicators were first asked to name up to 10 crops that they grew. They were then asked to name up to 15 pesticides used for each crop. This information was stored in an array of 10 variables for crops and 150 variables for the pesticides used on the crops. This means that the code for corn could be in any of 10 positions and that the pesticide name “atrazine” when associated with corn could be in any one of the 15 positions associated with the “corn” response in the 150-position array of pesticide names.

Although some less complicated array structures remain in the principal analysis files representing the CATI questionnaire responses,⁴ the data in the main arrays of crops, animals, and non-crops (e.g., fence row and storage area pesticide applications) have been moved to the “Private Applicator Supplemental Pesticide File,” the “Commercial Applicator Supplemental Pesticide File,” and the “Spouse Supplemental Pesticide File.” Each of these files has a separate record for each triple, consisting of a respondent ID, pesticide name, and the item to which the pesticide was applied. Each of these records includes the verbatim pesticide name as recorded in the questionnaire, a standardized version of that name, up to four Pesticide Chemical Codes (PC Codes)⁵ identifying the chemicals in the pesticide, the number of days in the year that the pesticide was applied, the average number of hours per day spent applying the pesticide, and several related variables. See the codebooks in the appendices for a detailed description of these verbatim pesticide files.

In the main Private Applicator and Spouse questionnaire files, a set of flags was created to indicate whether or not the respondent grew each of the crops listed in Question 12 of the questionnaires. Thus, the flag A2_CR_APPLES is set to 0 if the respondent did not grow apples and to 1 if he or she did grow apples. A corresponding set of numeric variables was created to store the values in the corresponding acreage question (Question 13). Thus the number of acres of apples grown is stored in the variable A2_CR_APPLES_ACRES. A similar set of flags was created for animals raised (Private Applicator Question 22, Spouse Question 18) and for noncrop applications (Private Application Question

⁴ These have all been converted to indicator variables as described later in this section.

⁵ A Pesticide Chemical Code (PC Code) is a six-digit number assigned by the Office of Pesticide Programs, U.S. Environmental Protection Agency, to identify pesticide chemicals. It is also called a Shaughnessy code. The PC Code is often used for searching computer databases because it is short and easy to enter.

28b, Spouse Question 25). A set of numeric variables corresponding to each animal records the number raised.

A set of flags similar to the crop flags was created in both the Private Applicator File, the Commercial Applicator File, and the Spouse File to identify chemicals that respondents used. Flags were created for all chemicals explicitly referenced in the Phase I questionnaire (50 chemicals)⁶ and an additional 33 chemicals targeted for research because they were either used by a large number of respondents or were suspected of having adverse health effects. We refer to this group of 82 chemicals as the “A list.” Flags were also created for any pesticide not on the A list that was used by 300 or more respondents (approximately 1% of the respondent applicators). We refer to this group as the “B list.” These flags can be used to determine whether or not someone was exposed to a chemical.

3.2.2.2 Variables in Supplemental Files Summarized in Analysis Files

The Private Applicator Supplemental Pesticide File, Commercial Applicator Supplemental Pesticide File, and the Spouse Supplemental Pesticide File contain detailed information on the number of days and hours the respondent applied each pesticide to each crop, animal, or noncrop to which he or she applied it. For each pesticide on the A list, the number of application days and the intensity-adjusted application days have been summarized at the pesticide level and placed on the respondent's record on the master Private Applicator File. Key Phase I exposure variables have also been placed on the Phase II master Private Applicator File and Phase II master Commercial Applicator File.⁷ An analysis file, the Private Applicator Base File, contains a core set of variables extracted from the master Private Applicator File (see Section 4.3.1).

For the pesticides that a person used in each application, the interviewer collected information on application methods. Information on use of personal protective equipment (PPE) was collected at the person level so that an adjusted exposure estimate could be made taking these practices into account. The adjustment was made by multiplying the total exposure in hours by an intensity score which was based on applicator behavior. The intensity algorithm used was comparable to that used during Phase I of the AHS (see M. Dosemici et al., 2002). An intensity score was calculated for each of the 82

⁶ The Phase I questionnaires have 50 pesticide variables. One of these is for permethrin and pyrethroid products applied to crops and one is for permethrin products applied in poultry, livestock, or animal confinement areas. They are treated separately because the application procedures and exposures are different for crops and animals.

⁷ The years of use and days of use variables were not collected for Spouses in Phase I.

chemicals on the A list when the data were available on the questionnaire. Since some of the chemicals on the list that were based on the Phase I questionnaire are no longer used, there was no exposure to them and thus no intensity scores were calculated for them. In Phase I, it was not possible to distinguish PPE used while mixing chemicals from that used in applying them. Both types of information were collected in Phase II. Phase II intensity scores use the separate data PPE data for mixing and applying chemicals and a combined PPE score when adjusting repair-related exposure estimates.

3.2.2.3 Functional and Chemical Class Flags

Two sets of related derived variables were also created and placed on the analysis files: chemical class flags and functional class flags. These are shown in Table 3–1 and Table 3–2. The chemical group flags have the same definitions as those on the Phase I data sets.

3.2.2.4 Conversion of Multiple Response Lists to Indicator Variables

Questions with lists that allow the interviewer to mark all the responses given by a respondent have been converted from lists of variables each of which could take any of the response codes to indicator variables for each possible response. Each indicator variable has the value of 0 if the applicator did not name the item or 1 if he or she did mention it.

These indicator variables are given a missing value if the respondent was not asked the question. If the respondent said “Don’t know” or refused to answer the question, a code of “.D” or “.R” was assigned. For example, Private Applicator Questionnaire Item 20a is:

20a. (During (Reference Year), which of the following types of nitrogen did you personally apply?

READ RESPONSES. MARK ALL THAT APPLY. PROBE FOR COMPLETE LIST. WHEN COMPLETE PRESS ENTER.

- | | |
|--|---------------------------------|
| 1. Anhydrous ammonia [Go to Q20b] | A2_ANHYD_AMMONIA |
| 2. Liquid nitrogen [Go to Q20b] | A2_LIQUID_NITROGEN |
| 3. Dry nitrate [Go to Q20b] | A2_DRY_NITRATE |
| 4. Regular or complete fertilizer [Go to Q20b] | A2_REGULAR_FERTILIZER |
| 5. Did not personally apply [Go to Q20b] | A2_NIT_DID_NOT_PERSONALLY_APPLY |
| 6. Something else... | A2_NIT_SOMETHING_ELSE |
- DK [Go to Q20b]
REF [Go to Q20b]

The names of the indicator variables created for each of the responses to item 20a are indicated to the right of the response. Thus, for example, the respondent indicate that he or she used regular fertilizer, the value for the variable A2_REGULAR_FERTILIZER would be set to 1. If the respondent did not indicate using regular fertilizer and did not refuse to answer the question or state “Don’t know,” then the value of A2_REGULAR_FERTILIZER was set to 0.

Table 3-1. Functional classes

Herbicides (Variables A2_RY_HRB, C2_RY_HRB, S2_RY_HRB) ⁸		
Alachlor products	EPTC products	Pendimethalin products
Atrazine products	Glyphosate products	Petroleum oil/petroleum distillates
Butylate products	Imazethapyr products	Trifluralin products
Chlorimuron ethyl products	Metolachlor products	2,4,5-T
Cyanazine products	Metribuzin products	2,4,5-TP (Silvex)
Dicamba products	Paraquat	2,4-D
Insecticides (Variables A2_RY_INS, C2_RY_INS, S2_RY_INS)		
Aldicarb products	Dichlorvos or DDVP products	Malathion
Aldrin	DDT	Parathion (ethyl or methyl)
Carbaryl products	Diazinon products	Permethrin or pyrethroid products
Carbofuran products	Dieldrin	Phorate products
Chlordane	Fonofos products	Terbufos products
Chlorpyrifos products	Heptachlor	Toxaphene
Coumaphos products	Lindane products	Trichlorfon products
Fungicides (Variables A2_RY_FNG, C2_RY_FNG, S2_RY_FNG)		
Benomyl products	Chlorothalonil products	Metalaxyl products
Captan products	Maneb or mancozeb products	Ziram products
Fumigants (Variables A2_RY_FUM, C2_RY_FUM, S2_RY_FUM)		
Aluminum Phosphide products	Ethylene dibromide products	
Carbon tetrachloride/carbon disulfide	Methyl bromide products	

⁸ The prefix "A2" is used in the Private Applicator File. The prefix "C2" is used in the Commercial Applicator File. The prefix "S2" is used in the Spouse File.

Table 3-2. Chemical classes

Organochlorine Insecticides (Variables A2_RY_OCH, C2_RY_OCH, S2_RY_OCH)		
Aldrin	Chlordane	DDT
Dieldrin	Heptachlor	Lindane
Toxaphene		
Organophosphate Insecticides (Variables A2_RY_OPH, C2_RY_OPH, S2_RY_OPH)		
Chlorpyrifos	Coumaphos	Diazinon
Dichlorvos/DDVP	Fonofos	Malathion
Parathion	Phorate	Terbufos
Trichlorfon		
Organothiophosphate Insecticides (Variables A2_RY_OTH, C2_RY_OTH, S2_RY_OTH)		
Chlorpyrifos	Coumaphos	Diazinon
Fonofos	Phorate	Terbufos
Carbamate Pesticides* (Variables A2_RY_CAR, C2_RY_CAR, S2_RY_CAR)		
Aldicarb	Carbofuran	Carbaryl
Benomyl		
Phenoxy Herbicides (Variables A2_RY_PNX, C2_RY_PNX, S2_RY_PNX)		
2,4-D	2,4,5-TP (Silvex)	2,4,5-T
Triazine Herbicides (Variables A2_RY_TRZ, C2_RY_TRZ, S2_RY_TRZ)		
Atrazine	Cyanazine	Metribuzin

* Includes insecticides and one fungicide.

Up to five responses were accepted during the interview. In the data releases in the series P2REL0312 (including P2REL0312.01 and P2REL0312.02) these responses were captured in the five variables:

- A2ONTYP1
- A2ONTYP2
- A2ONTYP3
- A2ONTYP4
- A2ONTYP5

Each one of these variables could have any of the six codes representing the six possible responses to the question as well the codes for Don't Know and Refused. To determine whether an applicator had applied liquid nitrogen using the original variables, it was necessary to search all five of them to see if the value "2" (liquid nitrogen) appeared in any of them. These variables have been removed from the dataset and replaced with the six indicator variables shown in Table 3-3, each having a value of 1 if the applicator indicated that he applied that type of nitrogen and zero (0) if the applicator answered the question but did not indicate using that type of nitrogen. These indicator variables are all set to ".D" if the respondent said "Don't know." They are all set to ".R" if the respondent refused to answer the question. They are set to missing if the respondent was not asked the question. Note that if the respondent said, "No" to the lead in question ("...was nitrogen in any other form applied to your farm land?"), question 20a was not asked and the indicator variables were set to missing.

Table 3-3. Indicator variable names associated with each Response Option to Question 20a in the Phase II Private Applicator Questionnaire

Response option	Variable name
1. Anhydrous ammonia	A2_ANHYD_AMMONIA
2. Liquid nitrogen	A2_LIQUID_NITROGEN
3. Dry nitrate	A2_DRY_NITRATE
4. Regular or complete fertilizer	A2_REGULAR_FERTILIZER
5. Did not personally apply	A2_NIT_DID_NOT_PERSONALLY_APPLY
6. Something else	A2_NIT_SOMETHING_ELSE

The other Private Applicator interview questions treated in this manner are those dealing with personal protective equipment (PPE) when mixing, loading or applying pesticides (Questions 35 and

36), a question regarding body parts exposed during a high pesticide exposure event (Question 39d), and foods excluded from the diet (Diet Module question 1a).

35. During (Reference Year), when you mixed or loaded pesticides did you normally wear any of the following special clothing or protective equipment:

1. Chemically resistant overalls like Tyvek
2. Chemically resistant boots/shoes (different than normal work boots)
3. Cartridge respirator, gas mask
4. Full face shield
5. Gloves
6. None

___DK

___REF

Table 3-4 shows the indicator variables corresponding to each response options in Question 35.

Table 3-4. Indicator variable names associated with each Response Option to Question 35 in the Phase II Private Applicator Questionnaire

Response option	Variable name
1. Chemically resistant overalls like Tyvek	A2_TYVEK_MIX
2. Chemically resistant boots/shoes (different than normal work boots)	A2_BOOTS_MIX
3. Cartridge respirator, gas mask	A2_RESP_MIX
4. Full face shield	A2_FACE_SHIELD_MIX
5. Gloves	A2_GLOVES_MIX
6. None	A2_NO_PPE_MIX

36. During (Reference Year), when you applied pesticides, did you normally wear any of the following special clothing or protective equipment?

1. Chemically resistant overalls like Tyvek
2. Chemically resistant boots/shoes (different than normal work boots)

3. Cartridge respirator, gas mask
4. Full face shield
5. Gloves
6. None

___DK

___REF

Table 3-5 shows the indicator variables corresponding to the response options for Question 36.

Table 3-5. Indicator variable names associated with each Response Option to Question 36 in the Phase II Private Applicator Questionnaire

Response option	Variable name
1. Chemically resistant overalls like Tyvek	A2_TYVEK_APPLY
2. Chemically resistant boots/shoes (different than normal work boots)	A2_BOOTS_APPLY
3. Cartridge respirator, gas mask	A2_RESP_APPLY
4. Full face shield	A2_FACE_SHIELD_APPLY
5. Gloves	A2_GLOVES_APPLY
6. None	A2_NO_PPE_APPLY

39d. (Regarding the most recent incident,) what parts of your body were exposed?

1. Head
2. Face (eyes, nose, mouth)
3. Torso, arms, legs
4. Hands
5. Feet
6. Lungs (from breathing)
7. Digestive tract (from swallowing)
8. Something else...

___DK

___REF

The variable names associated with each response to Question 39d are shown in Table 3-6.

Table 3-6. Indicator variable names associated with each Response Option to Question 39d in the Phase II Private Applicator Questionnaire

Response option	Variable name
1. Head	A2_HEAD
2. Face (eyes, nose, mouth)	A2_FACE
3. Torso, arms, legs	A2_TORSO_ARMS_LEGS
4. Hands	A2_HANDS
5. Feet	A2_FEET
6. Lungs (from breathing)	A2_LUNGS
7. Digestive tract (from swallowing)	A2_DIGESTIVE_TRACT
8. Something else...	A2_BP_SOMETHING_ELSE

Indicator variables were also created for question 1a from the Diet Module.

1a. Which of the following foods do you totally exclude from your diet?

1. Hamburgers and cheeseburgers
2. Beef steaks
3. Pork chops
4. Bacon
5. Sausage
6. Hot dogs or frankfurters
7. Chicken
8. Fish

The variable names associated with each response to Diet Module Question 1a are shown in Table 3-7.

Table 3-7. Indicator variable names associated with each Response Option to Diet Module Question 1a in the Phase II Private Applicator file

Response option	Variable name
2. Beef steaks	A2_EXCLUDE_BEEF
3. Pork chops	A2_EXCLUDE_PORK
4. Bacon	A2_EXCLUDE_BACON
5. Sausage	A2_EXCLUDE_SAUSAGE
6. Hot dogs or frankfurters	A2_EXCLUDE_HOTDOGS
7. Chicken	A2_EXCLUDE_CHICKEN
8. Fish	A2_EXCLUDE_FISH

The corresponding questions in the Commercial Applicator and Spouse Questionnaires were treated in the same way with the prefixes at the beginning of the variable names starting with “C2” for commercial applicators or “S2” for spouses.

3.2.2.5 Added Names of Active Ingredients Corresponding to the PC Codes

A set of four variables on the Phase II Supplemental Pesticide Files provide the names of the chemicals represented by the PC codes. AIName1 corresponds to PCCode1, AIName2 corresponds to PCCode2, AIName3 corresponds to PCCode3, and AIName4 corresponds to PCCode4. For example, Pursuit Plus contains two chemicals, Imazaquin (PC Code 128848) and Pendimethalin (PC Code 108501), so the PCCode and AIName variables have the values shown in Table 3-8.

Table 3-8. Values of PC code and active ingredient name variables for Pursuit Plus

Variable	Value
PCCode1	128848
PCCode2	108501
PCCode3	blank
PCCode4	blank
AIName1	IMAZAQUIN
AIName2	PENDIMETHALIN
AIName3	blank
AIName4	blank

3.2.2.6 Created Days per Year and Hours per Year Variables

Each record in the Phase II Private Applicator File, Commercial Applicator File, and Spouse File contains a set of days per year variables for each pesticide on the extended A list -- the list of 83 pesticides consisting of the 50 pesticides included in the Phase I questionnaire supplemented by 33 additional pesticides of high interest reported during Phase II. A corresponding set of 83 hours per year variables was also created.

The days per year variable for a specific respondent and specific pesticide was created by first determining the maximum number of days per year reported by the respondent for a pesticide for each application type (Crop/Animal/NonCrop). The days per year value for that pesticide for that person was calculated as the sum of these three maxima. That is, it was the sum of the maximum reported days per year over all crops for which that respondent used that pesticide, the maximum reported days per year over all animals for which that respondent used that pesticide, and the maximum reported days per year over all noncrops for which that respondent used that pesticide.

The hours per year variable for a specific respondent and specific pesticide was created by multiplying the reported days per year by the reported average hours per day for each reported instance of use of the pesticide and then summing these products over all reports across all application types.

3.2.3 Missing Value Codes

Missing character values are represented by blanks. Missing numeric values are set to “.D” if the respondent did not know the answer to a question and to “.R” if the respondent refused to answer a question. All other missing values, such as those caused by skips, are set to the standard SAS missing value of “.”.

3.2.4 Treatment of Differences among Versions of Modules

Differences among the different versions of the questionnaire modules required a certain amount of recoding of the earlier versions in order to make the responses commensurate with those in the final versions of the modules. The major differences and how they were treated are described below.

3.2.4.1 Differences between Versions 1 and 2 of Private Applicator Pesticide Use Module

There were a number of relatively minor wording differences between Versions 1 and 2 of the Private Applicator Pesticide Use module. For example, the stem for question 11 in Version 1 reads: “Since <reference date>, did you mix, load or apply pesticides for use on crops, animals, or for any other purpose?” In Version 2 of the module, the word “personally” was added before “mix, load or apply” so that the question reads: “Since <reference date>, did you personally mix, load or apply pesticides for use on crops, animals, or for any other purpose?” The following questions (with their corresponding variable names) had minor wording differences:

Q11 (A2MIXAPL)

Q11A (A2DMIXAPL)

Q15 (A2CPA)

Q16 (A2CPB)

Q17C (A2CPG)

Q20A (see Table 3-3 for variable names) – In addition to a small wording change, a new response (“Did not personally apply”) was added in Version 2. Thus no respondent who was interviewed using Version 1 has a value of 1 for A2_NIT-DID-NOT-PERSONALLY_APPLY.

Q20c (A2NTIMES)

In question 14 (“What insecticides, fungicides, fumigants, or herbicides (weed killers) were used on <crop> during <reference year>”), the interviewers were instructed in Version 1 to ask the respondents for the various chemicals that made up a manufacturer’s mixture. In later versions, they simply recorded the name of the manufacturer’s product. In order to make the responses easier to analyze, these responses were analyzed and up to four Pesticide Chemical Codes (PC Codes) were associated with each response. A PC Code is a six-digit number assigned by the Office of Pesticide Programs, U.S. Environmental Protection Agency, to identify pesticide chemicals.

3.2.4.2 Differences between Versions 2 and 3 of the Private Applicator Pesticide Use Module

Version 2 of the Private Applicator Pesticide Use module is very similar to Version 3, with some exceptions. The biggest difference is that Version 3 branches to a female health module, as appropriate.

The following questions were added to Version 3:

- Q1a. We currently have <applicator name> listed as <male or female inserted based on Phase I data>, is this correct? (A2VERGEN)
- Q3a. What was the month that you left (Iowa/North Carolina)? (A2LEFTSM)
- Q3b. What was the year that you left (Iowa/North Carolina)? (A2LEFTSY)
- Q54. Farming is more dangerous than jobs in industry or manufacturing. (A2DANGER)
- Q55. Accidents are just one of the occupational hazards of farming that must be accepted if you are going to be in the business. (A2ACCIDE)
- Q56. During a normal work week, it's common for me, while doing farm work, to experience a number of "close calls" that under different circumstances might have resulted in personal injury or property loss. (A2CLOSEC)
- Q57. To make a profit, most farmers take risks that might endanger their health. (A2PROFIT)

The number of verbatim responses to Question 24a1 which asked for a list of other pesticides used for animal applications was also expanded from three to six and the number of responses to Question 28c1, which asked for a list of pesticides used for noncrop applications, was expanded from 27 to 32. The number of pesticide application methods that could be recorded per crop in Question 14a was expanded from 3 to 15.

The following questions had minor wording differences:

- Q12. crops grown (e.g., A2_CR_APPLES)
- Q12A. What was the other nursery crop? (A2CONSP)
- Q15. In (reference year), how many days did you personally mix, load, or apply (pesticide) to (crop) (A2_DAYS)
- Q28b1. What was this [other non-crop application]? (A2NCAO)

Q32. In (reference year), how many days did you personally mix, load, or apply (pesticide) to (non-crop application) (A2_DAYS)

Question Q28b

In Versions 1 and 2, Question Q28b “During (reference year), what were these noncrop applications?” had nine responses. Two of these were combined in Version 3. The response choices in Version 2 were:

1. Weed or brush control for fence rows and other farm areas
2. Stored grains
3. Building applications
4. Nursery and garden applications (including shrubs, flowers and ornamentals)
5. Trees, including fruit trees
6. Lawn and turf applications
7. Rodent control
8. Grain bins/grain storage applications
9. Other, specify

The response choices in Version 3 were:

1. Weed or brush control for fence rows and other farm areas
2. Stored grains or grain bin applications
3. Building applications
4. Nursery and garden applications (including shrubs, flowers and ornamentals)
5. Trees, including fruit trees
6. Lawn and turf applications
7. Rodent control
8. Something else

Because farmers did not distinguish between stored grain applications (response 2 in the original version of the question) and grain bins/grain storage applications (response 8 in the original version), these two choices were combined in Version 3.

In follow-up questions to Question 28b, interviewers collected pesticide usage data for each noncrop application that a respondent performed. Ten of the 16 respondents who were administered Version 1 or Version 2 who indicated both choice 2 and choice 8 in the original question gave the same responses to all the related pesticide usage questions for those noncrop applications. Six respondents had differences in the pesticides that they listed for these two choices but failed to answer the related usage items such as number of days applied, average number of hours per day, and application method. In the Private Applicator File, responses to either choice 2 (stored grains) or the original choice 8 (Grain bins/grain storage applications) caused the indicator variable AN_NC_GRAINBIN to be set to 1 (Yes). In other words, for the purposes of setting the indicator variables, the response sets in Versions 1 and 2 for question 28b were mapped to the response set for Version 3. In the Private Applicator Supplemental Pesticide File; however, in order to preserve the fact that these respondents were exposed to the specific pesticides that they named, the file maintains the distinction in Versions 1 and 2 of the questionnaires even though this distinction was not made in the Version 3 interviews.

Question 31 (Variable A2NCMET)

The responses to Question 31 (Variable A2NCMET) were changed between Versions 2 and 3. In Version 2, the question was:

In <Reference Year>, what application method(s) did you use to apply the <pesticide> for <non-crop application>?

1. Hand-held sprayer (gun, wand, boom, or pump)
2. Broadcast from farm vehicle (NOT hand-held)
3. Backpack sprayer
4. Mist blower or fogger
5. Airblast
6. Banded spray
7. Gas canister
8. Dried pellets

9. Sealed packets
10. Dusting (including powders)
11. Something else...

In Version 3, the question was identical, but the response 10 dropped the parenthetical (including powders) and response “Spray Bottle” was added after response 10 and became response 11. The response “Something else” changed from a value of 11 to a value of 12.

1. Hand-held sprayer (gun, wand, boom, or pump)
2. Broadcast from farm vehicle (NOT hand-held)
3. Backpack sprayer
4. Mist blower or fogger
5. Airblast
6. Banded spray
7. Gas canister
8. Dried pellets
9. Sealed packets
10. Dusting
11. Spray bottle
12. Something else...

The response of “Something else” in Versions 1 and 2 was recoded from 11 to 12 so that the response values in the analysis files have the same meaning for all respondents.

Skip Patterns

There are minor differences in the skip patterns for the following variables:

Q17b (A2CPE)

Q12a (variable A2CONCSP) – “Nursery” in Version 2; “Other nursery” in Version 3

Q14 (variable A2_PESTICIDE_VERBATIM when asked for crop applications)

Q24 (variable A2_PESTICIDE_VERBATIM when asked for animal applications)

Q28 (variable A2_PESTICIDE_VERBATIM when asked for noncrop applications)

3.2.4.3 Differences among Versions of the Male Health Module

In Version 3 of the Male Health module, the response, “Yes for other reasons” was added to Q7b (Do you wear glasses to help correct nearsightedness or farsightedness or both?) (A2_GLASSES). The following variables were added in Version 3 of the Male Health module:

- Q7B1. Do you wear reading glasses? (A2_GLASSES_READING)
- Q7B2. Do you wear bifocals? (A2_GLASSES_BIFOCAL)

3.2.4.4 Differences among Versions of the Female Health Modules

In Version B of the Female Health modules, if menopause was natural and age less than or equal to 45, the questionnaire was designed to collect additional information. Version 2 of the module added question 21a., “Our records also indicate that this (menopause or end of menstrual periods) occurred at age (auto fill in from Q7 in Phase I Female and Family Health Questionnaire). Is this correct?” It was used when deciding which branch of the questionnaire to administer.

3.2.4.5 Differences among the Versions of the Cooking Practices Module

In Version 1 of the Cooking Practices module, descriptions of the various cooking methods were explicitly stated only once and then given on request. In later versions, they were given each time a cooking method question was asked.

In Versions 1 and 2 of the Cooking Practices module, the order of responses was different for several questions than in Version 3. The Version 1 and 2 responses for these questions have been recoded to be consistent with Version 3. The following questions were affected by the change:

- Q5c. (variable A2BASECF) What was the second most frequent way your bacon was cooked? – The fourth and fifth responses were interchanged (“Bought at a fast-food restaurant” and “Microwaved”).

- Q8a. (variable A2CKFRYO) About how often was the chicken you ate fried chicken or chicken nuggets? – The “Never” response was added in Version 3.
- Q11. (variable A2BBQCHA) During the past year, when you had grilled or barbecued meat, poultry, or fish, how often were they charred? – The order of the responses was reversed. Originally it went from Never (0%) to Almost always or Always (100%). In Version 3 it went from Almost always or Always (100%) to Never (0%).
- Q12. (variable A2BROWNE) During the past year, when you pan-fried or oven-broiled meat, poultry, or fish, how often were they well browned? – The order of the first five responses was reversed, similar to Q11. The sixth response (did not eat pan-fried or oven-broiled meats) stayed the same.

There are differences in the skip patterns for Cooking Practices Question 1a (variables A2TYPEE1 - A2TYPEE8).

3.3 Pesticide Exposure Flags

In the initial release of the Phase II data, an exposure indicator flag was set for each active ingredient on the extended A list and the B list. A value of 0 indicated that the applicator did not use the active ingredient. A value of 1 indicated that the applicator used the active ingredient. This indicator reflected responses to the question “What insecticides, fungicides, fumigants, or herbicides (weed killers) were used on [crop] during [Reference Year]?” and similar questions for animals and noncrops. If an applicator mentioned a pesticide, each active ingredient associated with the pesticide was given a flag value of “1.” If the applicator did not mention a pesticide containing a specific active ingredient, the flag was set to “0.”

A follow-up question for each pesticide named by the applicator asked, “In [Reference Year], how many days did you personally, mix, load, or apply [pesticide] to [crop]?” In some cases, people who had named a pesticide responded that they personally applied the pesticide for zero days or the days value was missing, i.e., they did not personally apply it. Only people who personally used the pesticide more than zero days per year had the ever/never used variable set to “1”. For example, Spouse variable S2_HERBICIDE_RY_A_1 indicated whether the respondent used Atrazine (PCCode=080803). A separate “used on farm” variable was created to indicate that the respondent reported use of the pesticide, whether or not he or she personally used it. This set of variables was named in a similar manner as the personal use indicator variables except that “RY” in the variable name was replaced with “UF”. For example, the Spouse variable for Atrazine is named S2_HERBICIDE_UF_A_1. Note that the “UF”

variables will always have a value of 1 when the corresponding “RY” variable has a value of 1; but the reverse is not true.

Note that because intensity adjusted hours calculations for each pesticide use responses to questions involving the use of personal protective equipment (PPE), mixing of chemicals, and repair of equipment, some respondents have pesticide exposure flags with values of “1” while having a corresponding intensity hour value of “missing.”

Note that when setting the value for the “RY” set of personal use variables, if the response to Private Applicator Question 11 is 0 (“No”) then we set the value of all 83 A-list ever/never used indicator variables and all 40 B-list ever/never used indicator variables to 0 rather than to missing. We did the same thing for the corresponding Commercial Applicator and Spouse Questionnaire items.

11. Since (date of enrollment), did you personally mix, load or apply pesticides for use on crops, animals, or for any other purpose?

- 1. Yes
- 0. No
- ___ REF
- ___ DK

If a respondent reported the use of any specific pesticides, but did not mention any of the specific A-list or B-list pesticides, the indicator variables for the non-mentioned A-list and B-list variables were set to 0. If, however, the respondent reported use of an unknown herbicide, an unknown insecticide, an unknown fumigant, or an unknown fungicide and reported personally mixing, loading, or applying that pesticide for at least one day, the values of the indicators and days of use for all A-list or B-list in that functional class of chemicals was set to missing unless the respondent explicitly indicated having used one of these pesticide. Note that if the respondent explicitly indicated having personally used a pesticide, the personal usage indicator for that pesticide was not changed, but remained set to 1.

Table 3-9 summarizes the changes in the values of one of these ever/never used variables.

Table 3–9. Change in values of methyl bromide indicator variable because of changes in definition of indicators

Value of indicator	Frequency in release P2REL0312	Frequency in release P2REL0506	Frequency in release P2REL0612
Missing	4124	0	1755
0	29154	29628	31544
1	179	22	158
2	Undefined	158	Undefined

Lifetime exposure indicators through Phase II for each A-List pesticide are derived from the corresponding indicator flags for Phase I exposure and Phase II exposure. If the respondent person indicated that he or she personally mixed, loaded, or applied a pesticide in either interview, the exposure indicator flag is set to 1 (“Yes”). If the respondent explicitly indicated no exposure in both interviews, the flag is set to 0 (“No”). If the response is missing in both interviews, the lifetime flag is set to missing. If the flag is missing in one interview and 0 (“No”) in the other, the lifetime flag is set to missing. This logic is summarized in Table 3–10

Table 3–10. Lifetime Indicator Variable through Phase II Logic

		P2 Indicator for Personally Applied Pesticide (For example: P2 Indicator for personally applied Atrazine = A2_HERBICIDE_RY_A_1): 0 = No; 1 = Yes		
		Missing	0	1
P1 Indicator for Personally Applied Pesticide (For example: P1 Indicator for personally applied Atrazine = A_HERBICIDE_CD1): 0 = No; 1 = Yes	Missing	Missing	Missing	1
	0	Missing	0	1
	1	1	1	1

3.4 Pesticide Exposure Algorithms

Exposure to pesticides during farming operations is not simply a matter of the amount of time a person is exposed to pesticides. A farmer spraying drops while riding in a tractor with an enclosed cab with an air filtration system is likely to receive a smaller effective dose of a pesticide per unit time than a farmer using a backpack sprayer to apply a pesticide to trees in an orchard. The intensity level of pesticide exposure is a function of pesticide handling procedures (i.e., mixing, loading, application and repairing of equipment) and protective equipment used. Algorithms for estimating exposure intensity are documented in Dosemeci et al. (2002).

Dosemeci et al. (2002) includes two algorithms for estimating the intensity of exposure. The first algorithm is based on questions in the Enrollment Questionnaire and it what is now termed Version 1 of the algorithm. The second algorithm used responses from both the Enrollment Questionnaire and the Take Home Questionnaire. Since only 40 percent of the applicators completed the Take Home Questionnaire, this version of the algorithm, which was called 'Algorithm 2' in the Phase I documentation, has not been used in analysis and is no longer included in AHS analysis files.

A new version of the intensity algorithm was defined in 2007 (Coble et al, forthcoming). This version uses the results of a sub-study of AHS applicators conducted by EPA that collected extensive biomarker data to calibrate the questionnaire responses (Thomas et al., in press). This 'Version 2' of the algorithm should be used in future analyses of AHS data.

The **intensity days** for Phase II was calculated by multiplying the average intensity score for all uses of the active ingredient for which an intensity score was calculated by the total days per year of use reported in Phase II for all uses of that active ingredient times the number years between the Phase II reference year and the Phase I reference year.

Lifetime days of personal use and **Lifetime Intensity Days** for each of the 50 pesticides from the enrollment questionnaire were calculated by summing the Phase I and Phase II days of use or the Phase I and Phase II Intensity Days, respectively. If days or intensity days from either Phase I or Phase II are missing, the corresponding lifetime days or **Lifetime Intensity Days** value was set equal to the non-missing value unless that value was zero. If the number of days of personal use for one phase was zero and the other was missing, the lifetime days of personal use for that pesticide was set to missing. If the

intensity days value for one phase was zero and the other was missing, the lifetime days or **Lifetime Intensity Days** for that pesticide was set to missing. The logic for setting the lifetime days of use variable (e.g., A2_HRB_A_LIFZE_DAYS1, Atrazine) is summarized in Table 3–11. The same logic was used for setting the Lifetime Intensity Days.

Table 3–11. Lifetime Exposure Days through Phase II Calculation

		Number of Days Reported in Phase II		
		(For example: P2 days reported for Atrazine times the number of years since the Phase I interview = A2_HRB_RY_A_TOTAL_DAYS_USE1)		
		Missing	0	>0
Number of Days Reported in Phase I	Missing	Missing	Missing	P2 Total Days
	0	Missing	0	P2 Total Days
	>0	P1 Days	P1 Days	P1 Days + P2 Total Days
(For example: Lifetime days thru Phase 1 using Atrazine = A_HERBICIDE_LIFE_DAY1)				

$$\text{Intensity Days} = \text{Total Application Days} * \text{Intensity Score}$$

Algorithm Intensity Score is a function of application methods used and PPE used. The Algorithm Intensity Score calculation is different in Version 1 and Version 2 of the algorithm.

For Version 1 of the algorithm, a combined PPE factor, which is used with the Repair Score, was defined as the minimum of the PPE_Mix factor and the PPE_Apply factor. The Algorithm Intensity Score is the sum of the product of each score and its corresponding factor.

$$\begin{aligned} \text{Algorithm Intensity Score} = & \text{[Mix Score]} * \text{[PPE_MIX]} + \\ & \text{[Apply Score]} * \text{[PPE_APPLY]} + \\ & \text{[Repair Score]} * \text{[PPE_COMBINED]} \end{aligned}$$

For Version 2, the Algorithm Intensity Scores for each application in the Phase II Supplemental File are calculated as follows:

Algorithm Intensity score = (Mix score + Apply score + Repair score) * PPE factor combined

Subjects with missing values for one or more of these four variables were assigned missing algorithm intensity scores. Subjects for whom the sum of the Mix Score, Apply Score, and Repair Score is zero will also be assigned missing algorithm intensity scores.

3.4.1 Average Intensity Score

In Phase II, the AHS applicators in some cases reported multiple use of the same active ingredient on different crops. If different application methods were used, then these separate applications of the same active ingredient may have different intensity scores. When a single applicator reported multiple uses of the same active ingredient with different intensity scores, a time-weighted average intensity score based on the reported hours per day across all these applications is calculated to obtain a single intensity score for each pesticide per subject.

3.4.2 Mix, Apply, and Repair Scores

Scores were created representing the pesticide-related activities of mixing pesticides, applying pesticides, and repairing equipment. For each pesticide applied to each crop, respondents were asked how many days they personally mixed, loaded or applied the pesticide to that crop.

15. In (Reference Year), how many days did you personally mix, load or apply (pesticide) to (crop)?
A2_DAYS

|_|_|_|

___DK [Go to Next Pesticide; At Last Pesticide Go to Q17]

___REF [Go to Next Pesticide; At Last Pesticide Go to Q17]

[IF Q15 = 0 THEN Go to Next Pesticide; At Last Pesticide Go to Q17]

If a respondent indicated that he or she applied mixed or applied the pesticide for zero days, then his or her mix and apply scores for this pesticide-crop combination was set to zero. If the respondent indicated a positive number of days for the pesticide-crop combination, then mix and apply scores were calculated as described below. Repair scores were determined independently of whether the respondent personally mixed or applied a pesticide.

Application of solid pesticide to crop using tractor	[a2cpo]
Application of gaseous pesticide to crop using tractor	[a2cpq]
Application of pesticide to crop not using tractor	[a2cps]
Application of <pesticide> to <animal>	[a2apapp]
Application of <pesticide> for <noncrop application>	[a2ncmet]

Each of the responses to these six questions has been mapped to a unique code that denotes the specific application method (a2_apply_method*). Where applicators reported more than one method for a particular crop and pesticide, the additional methods mentioned are coded as well. The apply score is assigned the appropriate weight for the application method using these codes. If more than one method is reported, a single application weighting factor was obtained by averaging the weights of each method.

18. (During (reference year), did you personally apply the (pesticide) to the (crop)?

- A2CPI 1. Yes
 0. No [If Q17a is no, DK. or Ref, Go to next pesticide for this crop]
 ___DK [If Q17a is no, DK. or Ref, Go to next pesticide for this crop]
 ___REF [If Q17a is no, DK. or Ref, Go to next pesticide for this crop]

23. Did you personally spray, dust or otherwise apply insecticides to your animals or to the animal confinement area in (Reference Year)?

- A2AINSEC 1. Yes
 0. No [Go to Q28]
 ___REF [Go to Q28]
 ___DK [Go to Q28]

30. (During (Reference Year)), did you personally apply (pesticide) for (non-crop application)?

- A2NCAPP 1. Yes
 0. No [Go to Q32]
 ___REF [Go to Q32]
 ___DK [Go to Q32]

Missing Application Method Scores

Subjects with missing application method information either did not personally apply, or were not asked about the application method. If the subject did not personally apply, the application score was set to 0 and the algorithm score calculated based on the mix, repair and PPE variables. If the subject indicated that they did personally apply, but the application method is missing, then the application method score was imputed as described in the section “Imputed Application Methods and Scores” below.

18. (During (reference year), did you personally apply the (pesticide) to the (crop)?

- A2CPI 1. Yes
 0. No [If Q17a is no, DK. or Ref, Go to next pesticide for this crop]
 ___DK [If Q17a is no, DK. or Ref, Go to next pesticide for this crop]
 ___REF [If Q17a is no, DK. or Ref, Go to next pesticide for this crop]

23. Did you personally spray, dust or otherwise apply insecticides to your animals or to the animal confinement area in (Reference Year)?

- A2AINSEC 1. Yes
 0. No [Go to Q28]
 ___REF [Go to Q28]
 ___DK [Go to Q28]

30. (During (Reference Year)), did you personally apply (pesticide) for (non-crop application)?

- A2NCAPP 1. Yes
 0. No [Go to Q32]
 ___REF [Go to Q32]
 ___DK [Go to Q32]

For subjects with a2_days > 0, but missing application method information, or apply method code =91:

If a2cpi=0, then the apply_score is set to 0, and the algorithm score is calculated if all other variables are non-missing.

If a2cpi=1 or missing, then the application method and score was imputed for herbicides, insecticides, fungicides, and fumigants as indicated in Table 3–12 (for private applicators), Table 3–13 (for spouses), or Table 3–14 (for commercial applicators), and the algorithm score was calculated using this imputed method score. For other pesticide types, the application method and method score was not imputed and the algorithm score was set to missing.

If a2ainsec=0, then the apply_score is set to 0, and the algorithm score is calculated if all other variables are non-missing.

If a2ainsec=1 or missing, then the application method and score was imputed for herbicides, insecticides, fungicides, and fumigants as indicated in Table 3–12 (for private applicators), Table 3–13 (for spouses), or Table 3–14 (for commercial applicators), and the algorithm score was calculated using

this imputed method score. For other pesticide types, the application method and method score was not imputed and the algorithm score was set to missing.

If a2ncappc=0, then the apply_score is set to 0, and the algorithm score is calculated if all other variables are non-missing

If a2ncapp=1 or missing, then the application method and score was imputed for herbicides, insecticides, fungicides, and fumigants as indicated in Table 3–12 (for private applicators), Table 3–13 (for spouses), or Table 3–14 (for commercial applicators), and the algorithm score was calculated using this imputed method score. For other pesticide types, the application method and method score was not imputed and the algorithm score was set to missing.

Imputed Application Methods and Scores

Missing application method scores for herbicides, insecticides, fungicides, and fumigants were imputed separately for each state (Iowa or North Carolina), participant type (private applicator, spouse, commercial applicator), pesticide type (herbicide, insecticide, fungicide, fumigant), and protection target (crop, animal, noncrop).

An imputed method name was assigned consisting of the characters “Imputed” followed by the participant code for private applicator (A2), spouse (S2), or commercial applicator (C2); the state (Iowa or NC); the pesticide type; and the target. Each element of the name was separated by an underscore to enhance readability. A sample imputed method name is:

Imputed_A2_Iowa_INSECTICIDE_CROP

An application method score for each of the imputed methods was calculated as the average application method score for the cell represent by the method for all observations with non-missing scores above zero.

When there are no observations for a cell in one state, the corresponding value from the other state was used. Therefore, since no Iowa spouses applied fumigants to crops, the application method value for North Carolina spouses applying fumigants to crops was used in both states.

Since the Commercial Applicator Questionnaire did not ask about the specific crops, animals, and other (noncrop) application for each pesticide, the protection target used for this imputation is constructed from the response to Question 14:

As a commercial pesticide worker, did you personally mix, load, handle or apply pesticides for:

1. Farm crops
2. Lawns or turf, including cemeteries and golf courses
3. Farm animals/shelters
4. Stored grain or other agricultural products
5. Interior or exterior building applications (home, commercial, or public)
6. Nursery or garden (including shrubs, flowers and ornamentals)
7. Weed or brush control (for fence rows, pastures, public lots, or highway right of ways)
8. Trees, including fruit trees
9. Rodent or termite control
10. Other applications

These were combined as follows:

Response Number	Protection Target
1, 2, 6, 8	CROP
3	ANIMAL
4, 5, 7, 9, 10	NONCROP

Table 4. Imputed Application Method Names and Scores for Private Applicators with Missing Application Methods.

Table 3–12. Imputed Application Method Names and Scores for Private Applicators with Missing Application Methods

Imputed Application Method	Imputed Score
Imputed_A2_Iowa_FUMIGANT_ANIMAL	65
Imputed_A2_Iowa_FUMIGANT_CROP	37
Imputed_A2_Iowa_FUMIGANT_NONCROP	66
Imputed_A2_Iowa_FUNGICIDE_ANIMAL	67
Imputed_A2_Iowa_FUNGICIDE_CROP	40
Imputed_A2_Iowa_FUNGICIDE_NONCROP	63
Imputed_A2_Iowa_HERBICIDE_ANIMAL	67
Imputed_A2_Iowa_HERBICIDE_CROP	39
Imputed_A2_Iowa_HERBICIDE_NONCROP	61
Imputed_A2_Iowa_INSECTICIDE_ANIMAL	64
Imputed_A2_Iowa_INSECTICIDE_CROP	28
Imputed_A2_Iowa_INSECTICIDE_NONCROP	70
Imputed_A2_NC_FUMIGANT_ANIMAL	66
Imputed_A2_NC_FUMIGANT_CROP	34
Imputed_A2_NC_FUMIGANT_NONCROP	47
Imputed_A2_NC_FUNGICIDE_ANIMAL	67
Imputed_A2_NC_FUNGICIDE_CROP	50
Imputed_A2_NC_FUNGICIDE_NONCROP	70
Imputed_A2_NC_HERBICIDE_ANIMAL	57
Imputed_A2_NC_HERBICIDE_CROP	41
Imputed_A2_NC_HERBICIDE_NONCROP	67
Imputed_A2_NC_INSECTICIDE_ANIMAL	53
Imputed_A2_NC_INSECTICIDE_CROP	40
Imputed_A2_NC_INSECTICIDE_NONCROP	71

Calculated using Private Applicator Supplemental Pesticide File: Release P2REL0712.00.

Table 3–13. Imputed Application Method Names and Scores for Spouses of Private Applicators with Missing Application Methods

Imputed Application Method	Imputed Score
Imputed_S2_Iowa_FUMIGANT_CROP	20
Imputed_S2_Iowa_FUMIGANT_ANIMAL	70
Imputed_S2_Iowa_FUMIGANT_NONCROP	64
Imputed_S2_Iowa_FUNGICIDE_ANIMAL	61
Imputed_S2_Iowa_FUNGICIDE_CROP	70
Imputed_S2_Iowa_FUNGICIDE_NONCROP	62
Imputed_S2_Iowa_HERBICIDE_ANIMAL	70
Imputed_S2_Iowa_HERBICIDE_CROP	47
Imputed_S2_Iowa_HERBICIDE_NONCROP	58
Imputed_S2_Iowa_INSECTICIDE_ANIMAL	66
Imputed_S2_Iowa_INSECTICIDE_CROP	60
Imputed_S2_Iowa_INSECTICIDE_NONCROP	73
Imputed_S2_NC_FUMIGANT_ANIMAL	70
Imputed_S2_NC_FUMIGANT_CROP	20
Imputed_S2_NC_FUMIGANT_NONCROP	53
Imputed_S2_NC_FUNGICIDE_ANIMAL	70
Imputed_S2_NC_FUNGICIDE_CROP	60
Imputed_S2_NC_FUNGICIDE_NONCROP	67
Imputed_S2_NC_HERBICIDE_ANIMAL	70
Imputed_S2_NC_HERBICIDE_CROP	50
Imputed_S2_NC_HERBICIDE_NONCROP	67
Imputed_S2_NC_INSECTICIDE_ANIMAL	58
Imputed_S2_NC_INSECTICIDE_CROP	62
Imputed_S2_NC_INSECTICIDE_NONCROP	73

Calculated using Spouse Supplemental Pesticide File: Release P2REL0712.00.

Imputed_S2_Iowa_FUMIGANT_CROP had no corresponding observations among Iowa spouses, so its value was set to be the same as Imputed_S2_NC_FUMIGANT_CROP.

Table 3–14. Imputed Application Method Names and Scores for Commercial Applicators with Missing Application Methods.

Imputed Application Method	Imputed Score
Imputed_C2_Iowa_FUMIGANT_CROP	38
Imputed_C2_Iowa_FUMIGANT_NONCROP	57
Imputed_C2_Iowa_FUNGICIDE_CROP	48
Imputed_C2_Iowa_FUNGICIDE_NONCROP	43
Imputed_C2_Iowa_HERBICIDE_CROP	43
Imputed_C2_Iowa_HERBICIDE_ANIMAL	60
Imputed_C2_Iowa_HERBICIDE_NONCROP	53
Imputed_C2_Iowa_INSECTICIDE_CROP	48
Imputed_C2_Iowa_INSECTICIDE_ANIMAL	68
Imputed_C2_Iowa_INSECTICIDE_NONCROP	63

Calculated using Commercial Applicator Supplemental Pesticide File: Release P2REL0712.00.

3.4.2.3 PPE Factors

During the Phase II interview, applicators were asked about PPE use when mixing/loading and when applying. Indicator values for each of the five types of PPE listed on the Phase II questionnaire were created based on the reported PPE. A parallel set of indicators and factors was created for PPE use during mixing bases on the responses to questions 36, 36a, and 36a1.

For Version 1 of the algorithm, if the response to either question was “yes,” then the value of the Mix Score was set to 9; otherwise the value was set to 0. The repair score for all respondents who reported repairs, regardless of whether they reported use of pesticides was set to 2.

Four PPE indicator variables for mixing operations were defined based on the responses to questions 35, 35a, and 35a1:

- PPE0_MIX – is 1 if no PPE was used during mixing and 0 if some PPE was used during mixing.

- PPE1_MIX – is 1 if chemically resistant boots or shoes or a full face shield were used during mixing and 0 if none of these items were used.
- PPE2_MIX – is 1 if chemically resistant overalls such as Tyvek or a cartridge respirator or gas mask were used and 0 if none of these items were used.
- PPE3_MIX – is 1 if chemically resistant, disposable, or waterproof gloves were used (responses 2, 3, or 4 to question 35a) and 0 if they were not used. PPE3_MIX was also set to 1 if the responses to question 35a1 indicated that the respondent used something equivalent to one of these responses.

The PPE indicators were then used to calculate the PPE factor for mixing or loading pesticides as follows.

If PPE0_MIX = 1, then A2_PPE_FACTOR_MIX = 1
 If PPE1_MIX = 1, then A2_PPE_FACTOR_MIX = 0.8
 If PPE2_MIX = 1, then A2_PPE_FACTOR_MIX = 0.7
 If PPE3_MIX = 1, then A2_PPE_FACTOR_MIX = 0.6
 If PPE1_MIX = 1 & PPE2_MIX = 1, then A2_PPE_FACTOR_MIX = 0.5
 If PPE1_MIX = 1 & PPE3_MIX = 1, then A2_PPE_FACTOR_MIX = 0.4
 If PPE2_MIX = 1 & PPE3_MIX = 1, then A2_PPE_FACTOR_MIX = 0.3
 If PPE1_MIX = 1 & PPE2_MIX = 1 & PPE3_MIX = 1, then A2_PPE_FACTOR_MIX = 0.1

The resulting PPE factor is named A2_PPE_FACTOR_APPLY. An additional variable, A2_PPE_FACTOR_COMBINED is the minimum of the MIX and APPLY PPE factors.

For Version 2 of the algorithm, the applicators who reported wearing chemically resistant gloves (a2_glove_type_mix/apply=2, 3, or 4) will get a 60% reduction in intensity. An additional 10% reduction was assigned for each additional type of PPE (denoted by the indicator variables a2_tyvek_mix/apply, a2_boots_mix/apply, a2_resp_mix/apply, and a2_face_shield_mix/apply) up to a limit of 3 additional items. The PPE factor was calculated as follows.

The PPE factor was calculated using the formula:

$$\text{PPE factor} = 1 - (0.6 * \text{CR glove} + 0.1 * \min(\text{sum}(\text{tyvek} + \text{boot} + \text{resp} + \text{face}), 3))$$

Two PPE factors were calculated, one for mixing and one for applying. The minimum (i.e., more protective) of the two values was used for calculation of the Phase II intensity score.

$$\text{PPE factor combined} = \min(\text{a2_ppe_factor_mix}, \text{a2_ppe_factor_apply})$$

3.4.2.4 Repair Score

For version 1, if the respondent reported repairing equipment, but not to pesticide use, a score of 2 was assigned. In version 2, the repair score is assigned a value of 20 if applicators reported performing repairs themselves during the past year (A2REPAIR = 1). For both versions, if they did not report repairing equipment in the last year then they are assigned a score of 0 (A2REPAIR = 0).

If the A2REPAIR variable is missing, then the A2_REPAIR_SCORE variable was set to missing.

3.5 Version 2 General Considerations

When creating new algorithm-related variable names, the same names were used as previously, but the characters “V2” were added to the variable name. For example, A2_MIX_SCORE became A2_V2_MIX_SCORE.

Table 3-15. Assignment of Method Codes to Other Specify Responses for Each Question

Variable Name	Other Method	Method_1	Method_2	Method_3
A2APAO	AEROSOL CAN	11		
A2APAO	APPLICATOR GUN	38		
A2APAO	APPLIED BY HAND IN THE RAFTORS	27		
A2APAO	APPLIED DUST BY HAND ON BACK	38		
A2APAO	APPLIED ORALLY	33		
A2CPN	1/2 BANDED 1/2 BROADCAST	16	18	
A2CPN	1/2 BANED-- 1/2 BROADCAST SPRA	16	18	
A2CPN	16FT BOOM SPRAY	17		
A2CPN	ON A FINISHER-(BROADCAST/INCOR	32		
A2CPN	ONE THIRD BANDED AND TWO THIRD	16	18	
A2CPP	GRANULAR, INSECTICIDE HOPPER	15		
A2CPP	GRANULATED	15		
A2CPP	HALF BROADCAST , HALF BANDED	16	18	
A2CPP	HALF BROADCAST AND HALF BANDED	16	18	
A2CPR	ENCLOSED CHEMICAL HANDLING SYS	91		
A2CPR	INCORPORATEDIT	32		
A2CPR	LAID ON TOP	18		
A2CPR	SPRAYED TOP OF SOIL&CULTIVATED	32		
A2CPT	APPLIED A TEASPOON FULL TO TOP	91		
A2CPT	APPLIED A TEASPOONFULL AT TOP	91		
A2CPT	SPOONED ONTO TOP OF SOIL	91		
A2CPT	SPREAD BY HAND...	91		
A2CPT	SPRINKLES IT ON...	91		
A2CPT	SPIURT BOTTLE	27		
A2CPT	SQUEEZE BOTTLE	27		
A2NCMOT	BAR FORM PLACED BY HAND IN AREA	14		
A2NCMOT	BAR IN BUCKET-PUT OUT AS NEEDED	46		
A2NCMOT	BAR-BAIT	14		
A2NCMOT	BARS	14		
A2NCMOT	BARS OF SIX TO PACKAGE	14		
A2NCMOT	BATTERY SPRAYER	18		
A2NCMOT	BLOCK LAID OUT	14		

Table 3-16. Application Method Scores Associated with Each Application Method Code

Method Code	METHODS	Old Score (Used in calculations in Release P2REL0612.03)	New Score
10	Aerial	1	10
11	AerosolCan	7	70
12	Airblast	9	90
13	BackpackSprayer	8	80
14	BaitBars_Blocks_Cubes	1	10
15	BandedApplication	2	20
16	BandedSpray	2	30
17	BoomOnTractor_Truck_OrTrailer	3	40
18	BroadcastApplication	3	40
19	DipAnimals	5	50
20	DistributableTablets_Granules_Pellets	1	10
21	DustAnimals	7	70
22	EarTag	1	10
23	Fog_Mist_Animals	9	90
24	GardenHose	7	70
25	GasCanister	2	20
26	GreenhouseSprayer	9	70
27	Hand_HeldSqueezeOrSquirtBottle	7	70
28	HandSpray	9	70
29	HandSpreaderOrPushSpreader	2	20
30	HangPestStripsInBarn	2	20
31	In_Furrow	2	20
32	Incorporated	2	20
33	InjectAnimals	2	20
34	MistBlowerOrFogger	9	80
35	PersonallyAppliedToSeed	2	20
36	PlanterBox	2	20
37	PourFumigantFromBucket	9	90
38	PourOnAnimals	7	70
39	PowderDuster	9	90
40	Pre_TreatedSeed	1	10
41	SealedPackets	1	10
42	SeedTreatment	1	10

Method Code	METHODS	Old Score (Used in calculations in Release P2REL0612.03)	New Score
43	SoilInjectedOrDrilled	4	40
44	SprayAnimals	6	60
45	SprayOverRows	4	40
46	WateringCanOrSprinklingCan	7	70
47	Appliedwithbrushspooncupetc		70
48	Spreadbyhand_handbroadcastgranules		40
91	Other		

Table 3-17. Phase II PPE Variables:

<u>variable</u>	<u>content</u>
a2_tyvek_mix	p35: used chemically resistant overalls like Tyvek when mixing or loading pesticides
a2_boots_mix	p35: wore chemically resistant boots/shoes when mixing or loading
a2_resp_mix	p35: used respirator when mixing or loading pesticides
a2_face_shield_mix	p35: wore full face shield when mixing or loading pesticides
a2_gloves_mix	p35: wore gloves when mixing or loading pesticides
a2_ppe_none_mix	did not use PPE when mixing or loading pesticides
a2_glove_type_mix	p35a: glove type worn when mixing pesticides (code)
a2_glove_type_mix_text	glove type worn when mixing pesticides (description)
a2_boots_apply	p36: wore chemically resistant boots/shoes when applying pesticides
a2_resp_apply	p36: used respirator when applying pesticides
a2_face_shield_apply	p36: wore full face shield when applying pesticides
a2_gloves_apply	p36: wore gloves when applying pesticides
a2_ppe_none_apply	did not use PPE when applying pesticides
a2_glove_type_apply	p36a: glove type worn when applying pesticides (code)
a2_glove_type_apply_text	glove type worn when applying pesticides (description)
a2_ppe0_mix	indicator flag for no use of PPE when mixing or loading pesticides
a2_ppe1_mix	indicator flag for use of chemically resistant boots, face shield, or fabric or
a2_ppe2_mix	indicator flag for use of chemically resistant coveralls or respirator when mixing
a2_ppe3_mix	indicator flag for use of chemically resistant, rubber, or disposable gloves when mixing
a2_ppe0_apply	indicator flag for no use of PPE when applying pesticides
a2_ppe1_apply	indicator flag for use of chemically resistant boots, face shield, or fabric or
a2_ppe2_apply	indicator flag for use of chemically resistant coveralls or respirator when applying
a2_ppe3_apply	indicator flag for use of chemically resistant, rubber, or disposable gloves when applying
a2_ppe_factor_mix	PPE factor for mixing pesticides
a2_ppe_factor_apply	PPE factor for applying pesticides
a2_ppe_factor_combined	PPE factor - combined for mixing, loading, and applying pesticides
a2_mixppe	use PPE when mixing or loading pesticides
a2_applyppe	use PPE when applying pesticides

Table 3–18. New Application Method Codes and Scores

New Code	METHODS ON PHASE QX III CATI LIST	Old Score	New Score
10	10. Aerial	1	10
11	11. AerosolCan	none	70
12	12. Airblast	9	140
13	13. BackpackSprayer	8	80
14	14. BaitBars_Blocks_Cubes	none	10
15	15. BandedApplication (granular)	2	20
16	16. BandedSpray	2	30
17	17. BoomOnTractor_Truck_OrTrailer	3	40
18	18. BroadcastApplication	3	40
19	19. DipAnimals	5	50
20	20. DistributableTablets_Granules_Pellets	1	10
21	21. DustAnimals	7	70
22	22. EarTag	1	10
23	23. Fog_Mist_Animals	9	90
24	24. GardenHose	none	70
25	25. GasCanister	2	20
26	26. GreenhouseSprayer	9	70
27	27. Hand_HeldSqueezeOrSquirtBottle	none	30
28	28. HandSpray	9	70
29	29. HandSpreaderOrPushSpreader	none	20
30	30. HangPestStripsInBarn	2	20
31	31. In_Furrow	2	20
32	32. Incorporated	2	20
33	33. InjectAnimals	2	20
34	34. MistBlowerOrFogger	9	80
35	35. PersonallyAppliedToSeed	none	40
36	36. PlanterBox	none	20
37	37. PourFumigantFromBucket	9	90
38	38. PourOnAnimals	7	70
39	39. PowderDuster	9	90
40	40. Pre_TreatedSeed	1	20
41	41. SealedPackets	1	10
42	42. SeedTreatment	1	40
43	43. SoilInjectedOrDrilled	4	40
44	44. SprayAnimals	6	70
45	45. SprayOverRows	none	30
46	46. WateringCanOrSprinklingCan	none	30
91	91. Other		?
98	98. Did Not Apply		

Table 3–19. Recode of Phase II Application Method using P3 Codes

Variable	Response	Phase 3 Method Code
a2apapp	1	21
a2apapp	2	19
a2apapp	3	33
a2apapp	4	22
a2apapp	5	91
a2cpl	1	18
a2cpl	2	16
a2cpl	3	28
a2cpl	4	34
a2cpl	5	91
a2cpo	1	18
a2cpo	2	16
a2cpo	3	91
a2cpq	1	91
a2cpq	2	91
a2cps	1	28
a2cps	2	13
a2cps	3	16
a2cps	4	25
a2cps	5	34
a2cps	6	42
a2cps	7	42
a2cps	8	26
a2cps	9	39
a2cps	10	91
a2ncmot	1	28
a2ncmot	2	18
a2ncmot	3	13
a2ncmot	4	34
a2ncmot	5	12
a2ncmot	6	16
a2ncmot	7	25
a2ncmot	8	20
a2ncmot	9	91
a2ncmot	10	39
a2ncmot	11	27
a2ncmot	12	91

Body Mass Index Estimates

Where sufficient information was available, estimates were made of each participant's body mass index (BMI) at enrollment based on responses to the Phase I questionnaire about height and weight supplemented by Phase II height and weight questions and some external information. The estimation process addressed a number of data quality issues including both within-questionnaire issues and inconsistency across sources of information. The estimation methods are described in Appendix C. The resulting BMI estimates have been included in the Demographic File.

4. DATA FILE DESCRIPTIONS

The Phase II data consist of a set of related files which can be linked to each other and to the Phase I files by the study IDs of the respondents. The following files are part of the Phase II data:

- Private Applicator File
- Private Applicator Supplemental Pesticide File
- Spouse File
- Spouse Supplemental Pesticide File
- Commercial Applicator File
- Commercial Applicator Supplemental Pesticide File
- Diet History Questionnaire File

The appendices contain a codebook for each of these files providing a detailed description of its contents. Each of the Phase II files is discussed briefly in a separate section below.

4.1 Private Applicator File

The Private Applicator File contains the responses to the questions in the Phase II CATI interviews of private pesticide applicators. It contains 33,457 records. Of these; 21,089 are from Iowa farmers and 12,368 are from North Carolina farmers. It also contains a number of derived variables that will be of use to analysts. A number of these record administrative information. These are summarized in Table 4-1.

Table 4-1. Description of administrative variables

Administrative Variable	Description
A2_VERSION (P1, P2, P3)	Questionnaire version
A2_STATE (IA, NC)	Field station administering interview
A2_REFERENCE_YEAR	Reference year, calculated from the variables A2ACFARM, A2STFARM, A2LAFARM, and A2_SASDATE_STARTDA
A2_DECBOX	Path coming out of the Decision Box (before the Crop section):

	Path = 1 --> Go through all modules and all questions within each module
	Path = 2 --> Go through all modules but not all questions within a module
	Path = 3 --> Skip the Crop and Animal modules, go straight to the NonCrop module
	Path = 4 --> Just do the Health and Diet modules, all other modules are skipped
A2_HEALTHMOD	Health module given
A2_FS_COMPLETE	Indicator for complete interview (from field station files)
A2_MOND_COMPLETE	Indicator for complete interview (calculated from module completion flags)
A2_START_MODFLAG	Start Module completion flag ⁹
A2_HEALTH_MODFLAG	Health Module completion flag
A2_CROP_MODFLAG	Crop Module completion flag
A2_ANIMAL_MODFLAG	Animal Module completion flag
A2_NONCROP_MODFLAG	Noncrop Module completion flag
A2_HIEXP_MODFLAG	High Exposure Module completion flag
A2_DIET_MODFLAG	Diet Module completion flag
A2_FILL_LOAD	Indicator for personally fill or load application equipment for crops and noncrops
A2_PERSONALLY_APPLY	Indicator for personally apply for crops, animals, and noncrops
A2_FARM_VEHICLE	Indicator for Q18a = Yes for any crop pesticide or Q31 = 2 for any NonCrop pesticide

Note that while most of the questions in the CATI questionnaire are directly represented in the file by a single variable, the crop, animal and noncrop sections of the Pesticide Use module have been restructured to make them easier to use in analyses. In particular, indicator variables were created to reflect crops, animals, and noncrops to which pesticides were applied and to reflect which chemicals were used (see Section 3.2.2). Details about specific applications of pesticides have been placed in the Private Applicator Supplemental Pesticide File (see Section 4.2).

The questionnaire includes a number of items that record dates. These include preloaded dates such as the Phase I enrollment date. It also includes a number of dates provided as answers to questions. Most of these are full dates but some consist only of month and year. All date responses consist of separate variables for month, day (when asked), and year. Year variables are given in both 2-digit and 4-digit format. When the questionnaire requested the full date, a SAS date variable is also included. Thus, for example, the Phase II interview start date is represented by the following five variables:

⁹ Module completion flags have three possible values: 0 = module not begun; 1 = module begun, but not completed; 2 = module completed.

A2_START_DD	Day on which the interview was started
A2_START_MM	Month in which the interview was started
A2_START_YY	Year in which the interview was started (2 digit)
A2_START_YYYY	Year in which the interview was started (4 digit)
A2_SASDATE_STARTDA	Full SAS date

Indicator variables were created for each crop, animal, and noncrop and the pesticides listed in Tables 3-1 and 3-2. These indicators have the value 1 if the respondent grew the crop or applied the pesticide and 0 if he or she did not do so. There is a separate indicator variable for each of the crops named in the questionnaire. For example, the indicator variables for apples, alfalfa, and barley are A2_CR_APPLES, A2_CR_ALFALFA, and A2_CR_BARLEY, respectively.

It was not useful to create indicator variables for all of the thousands of pesticides named by the respondents. A pesticide was assigned an indicator variable if it was one of the 49 chemicals listed in the Phase I questionnaires, if it was suspected of having adverse health effects, or if it was used by 300 or more applicators. Because there were many potential crop-pesticide, animal-pesticide, and noncrop-pesticide pairs, more detailed information about pesticides and how they were used was placed in a separate analysis file called the Private Applicator Supplemental Pesticide File, which is described in the next section.

The Private Applicator File contains the indicator variables for the chemicals on the A list and the B list (see Section 3.3).¹⁰ It also contains indicator variables for the functional and chemical groups with the same list of chemicals as defined in Phase I. The Phase II variable names contain the mnemonic “RY” for reference year to emphasize that these flags refer only to exposures during the reference year asked about during the CATI interview. Table 4-2 lists the variable names for these indicators.

Table 4–2. Indicator variables for functional and chemical class groups

Indicator Variable Name	Description
A2_RY_FNG	Pesticide usage/Fungicides

¹⁰See the appendices for the names of chemicals in the A and B lists.

A2_RY_FUM	Pesticide usage/Fumigants
A2_RY_HRB	Pesticide usage/Herbicides
A2_RY_INS	Pesticide usage/Insecticides
A2_RY_CAR	Pesticide usage/Carbamates
A2_RY_OCH	Pesticide usage/Organochlorines
A2_RY_OPH	Pesticide usage/Organophosphates
A2_RY_OTH	Pesticide usage/Organothiophosphates
A2_RY_PNX	Pesticide usage/Phenoxy
A2_RY_TRZ	Pesticide usage/Triazine

The master Private Applicator File contains the following set of fields for each of the 82 specific pesticides on the A list, as well as the Phase II indicator flags for the pesticides on the B list.

1. Phase II Personal Use Indicator Flag (e.g., A2_FUMIGANT_RY_A_1, which is the indicator flag for Methyl Bromide (PCCode=053201)). --These flags take the values:

0 = did not report *personal* use of this pesticide
1 = reported *personal* use of this pesticide
Missing = not in supplemental file¹¹
2. Phase II Used on Farm Indicator Flag (e.g., A2_FUMIGANT_UF_A_1, which is the indicator flag for Methyl Bromide (PCCode=053201)). --These flags take the values:

0 = did not report *any* use of this pesticide
1 = reported use of this pesticide (whether or not any personal use was reported)
Missing = not in supplemental file⁹
3. Phase II Days per Year of Use (e.g., A2_FUM_RY_A_DAYSPERYR_1) – These variables are the **sum** of the maximum number of days pesticides were applied to **crops**, plus the maximum number of days pesticides were applied to **animals**, plus the maximum number of days pesticides were applied to **noncrops** (e.g., fence rows).
4. Total days of use since enrollment (e.g., A2_FUM_RY_A_TOTAL_DAY_USE1) – These variables are the product of the number of Phase II years times the number of Phase II Days per Year. Phase II years are calculated as the difference between the Phase II reference year and the enrollment year.
5. Average algorithm intensity score for this pesticide calculated from Phase II Supplemental file (e.g., A2_FUM_RY_A_AVG_ALG_INTSCORE1).
6. Phase II Intensity Days (e.g., A2_FUM_RY_A_AVG_ALG_INTENSITYDAYS_1) – These variables are equal to the total days of use since enrollment times the average algorithm intensity score.

The master Private Applicator File also contains combined Phase I and Phase II pesticide exposure data for each of the pesticides on the A-list:

7. Cumulative Lifetime Days (e.g., A2_FUM_RY_A_LIFE_DAYS1) – The sum of the Phase I Days plus the Phase II Days for the pesticide.
8. Cumulative Lifetime Intensity Days (e.g., A2_FUM_RY_A_LIFE_INTDAYS1) – The sum of the Phase I Intensity Days and the Phase II Intensity Days for the pesticide.

¹¹ If the respondent reported at least one day of use of an unknown herbicide, an unknown insecticide, an unknown fumigant, or an unknown fungicide, the values of the all A-list or B-list indicators in that functional class of chemicals were set to missing unless the respondent explicitly indicated having used a specific pesticide. In that case the indicator variable for the explicitly mentioned pesticide was set to 1.

The file also contains a number of demographic variables that are identical to those in the Phase I file. They are listed in Table 4-3.

Table 4-3. Phase I demographic variables included in Phase II Private Applicator file

Phase I variable	Description
SPSPAIR	Identifies the household ID, if any, that was assigned to the spouse of an applicator who also enrolled as an applicator
ARACE	Race
AHISPAN	Hispanic origin flag
ASCHOOL	Highest level of schooling
AGE_AT_ENROLLMENT	This variable has the same values as the Phase I variable A_AGE and represents the respondent's age at the time he or she completed the Enrollment Questionnaire

The exposure estimate variables from the Phase I Private Applicator File are also provided in the Phase II Private Applicator File.

4.2 Private Applicator Supplemental Pesticide File

The Private Applicator Supplemental Pesticide File contains one record for each triple of a participant ID, crop/animal/noncrop to which a pesticide was applied, and pesticide name. There are 190,015 records in the file.

The file contains the pesticide name as recorded by the interviewer; a standardized pesticide name; the specific crop, animal, or noncrop to which the pesticide was applied by the respondent; an indicator variable identifying the target usage as crop, animal, or noncrop; an indicator variable identifying the pesticide type as herbicide, insecticide, fumigant, or fungicide; PC codes for each pesticide (any not needed are set to missing); a flag indicating for crops only whether the pesticide was one of the “top four” used by the respondent (as determined by the number of days it was applied); and the detailed information asked about pesticides applied to crops, animals, and noncrops.

4.3 Spouse File

The Spouse File contains responses to the questions in the Phase II CATI interviews of spouses of private applicators. It contains 23,796 records. Of these, 16,931 are from Iowa spouses and 6,865 are from North Carolina spouses. The general structure of the file is similar to that for the Private Applicator File. Refer to the Spouse Codebook in the appendices for a detailed description of the file.

The administrative variables for the Spouse File are parallel to those for the Private Applicator file. Because of the more complicated branching structure in the Spouse Questionnaire, however, there are eight paths, determined by the decision box. Thus the Spouse File decision box variable, S2_DECBOX, has the following meanings:

Path = 1: Go through all modules

Path = 2: Go through all modules except the NonCrop module

Path = 3: Go through all modules except the Animal module

Path = 4: Go through all modules except the Animal and NonCrop modules

Path = 5: Go through all modules except the Crop module

Path = 6: Go through all modules except the Crop and NonCrop modules

Path = 7: Go through all modules except the Crop and Animal modules

Path = 8: Go through all modules except the Crop, Animal, and NonCrop modules

Note that, like the Private Applicator File, the Spouse File contains demographic variables copied from the Phase I Spouse File. The variable names for the demographic variables in the Spouse File are shown in Table 4-4.

Table 4-4. Phase I demographic variables included in Phase II Spouse File

Phase I Variable	Description
SPSPAIR	Identifies spouse pair ID
SRACE	Race
SHISPAN	Hispanic origin flag
SSCHOOL	Highest level of schooling

Note that, because years and days of applying specific pesticides were not collected for spouses in Phase I, the only exposure variables copied from the Phase I Spouse File are those indicating whether a respondent was or was not exposed to each of the pesticides listed in Phase I Spouse Questionnaire.

4.4 Spouse Supplemental Pesticide File

The Spouse Supplemental Pesticide File contains one record for each triple of a participant ID, crop/animal/noncrop to which a pesticide was applied, and pesticide name. There are 26,852 records in the file. The structure of the file is the same as the structure of the Private Applicator Supplemental Pesticide File. Refer to the Spouse Supplemental Pesticide Codebook in the appendices for a detailed description of the file.

4.5 Commercial Applicator File

The Commercial Applicator File contains the responses to the questions in the Phase II CATI interviews of commercial pesticide applicators. It contains 2,885 records. It also contains a number of derived variables that will be of use to analysts. A number of these record administrative information.

4.6 Commercial Applicator Supplemental Pesticide File

The Commercial Applicator Supplemental Pesticide File contains one record for each triple of a participant ID, type of application for which a pesticide was applied, and pesticide name. There are 22,924 records in the file.

The file contains the pesticide name as recorded by the interviewer; a standardized pesticide name; the specific way in which the pesticide was used by the respondent (response to Q1`4); an indicator variable identifying the pesticide type as herbicide, insecticide, fumigant, or fungicide; PC codes for each pesticide (any not needed are set to missing); a flag indicating for crops only whether the pesticide was one of the “top four” used by the respondent (as determined by the number of days it was applied); and the detailed information asked about pesticides applied to crops, animals, and noncrops.

4.7 Diet History Questionnaire File

The Diet History Questionnaire File contains one record for each Diet History Questionnaire that was returned to the project. There are a total of 35,011 records in the file. Of these, 26,228 are from Iowa participants and 8,783 are from North Carolina participants. There are no duplicate IDs in the file.

5. USAGE NOTES

The purpose of this chapter is to provide guidance to users with respect to issues that they are likely to encounter while analyzing the Agricultural Health Study data files. The following topics are included:

- How to link files for analyses;
- Discrepancies between Phase I and Phase II Files;
- Definition of Exposure Measures;
- Unusual Values;
- Interpretation of Missing Data Patterns; and
- Topics Discussed in Phase I User Manual that May Be of Interest.

Each of these topics is discussed in a separate section below.

5.1 How to Link Files for Analyses

All Phase I and Phase II AHS files contain a key variable named PARTID, which is the household identifier assigned at enrollment to a pesticide applicator. The value for PARTID is identical in the Private Applicator Files and Spouse Files for husbands and wives. Since the questionnaire files contain either applicator data or spouse data, PARTID is sufficient to uniquely identify a participant within one of these files.

Several files contain data from both applicators and spouses. These are the Demographic File, the Cancer Registry File, and the Mortality File. The Demographic File contains one record for each

member of the AHS cohort. To correctly match records in the Demographic File to one of the questionnaire files, it is necessary to use the variable AP_SPOUSE in conjunction with PARTID. AP_SPOUSE has the value “C” for commercial applicators, “P” for private applicators (farmers), and “S” for spouses of private applicators.

The Cancer Registry File and the Mortality File also contain records for both applicators and spouses. To correctly match them to the questionnaire data files, it is necessary to use either the AP_SPOUSE variable or (for the Cancer Registry File) the gender variable in each file in conjunction with PARTID. The gender variable in all AHS analysis files has the value 1 for males and 2 for females. The availability of the AP_SPOUSE variable and the availability and name of the gender variable in each of the main AHS files is summarized in Table 5-1. Note that there are males and females in both the Private Applicator File and the Spouse File.

Table 5-1. Gender and Applicator/Spouse Variable Names and Availability in AHS Data Files

File	File Type	Gender variable	Applicator/Spouse variable
Phase I Private Applicator	Master	AGENDER	N/A
Phase I Commercial Applicator	Master	AGENDER	N/A
Phase I Spouse	Master	SGENDER	N/A
Phase II Private Applicator	Master	A2GENDER	N/A
Phase II Private Applicator Supplemental Pesticide	Master	A2GENDER	N/A
Private Applicator Base	Analysis	A2GENDER	N/A
Private Applicator Pesticide Usage	Analysis	A2GENDER	N/A
Private Applicator Diet and Health	Analysis	A2GENDER	N/A
Phase II Spouse	Master	S2GENDER	N/A
Phase II Spouse Supplemental Pesticide	Master	S2GENDER	N/A
Demographic	Master	GENDER	AP_SPOUSE
Cancer Registry	Master	GENDER	AP_SPOUSE
Mortality	Master	N/A	AP_SPOUSE
Buccal Cell Inventory	Master	N/A	AP_SPOUSE
Diet History Questionnaire	Master	SEX	N/A

5.2 Discrepancies between Phase I and Phase II Files

Analysts should be aware of a number of discrepancies between similar data items in the Phase I and Phase II data files. These are most apparent with Phase I values preloaded into the CATI questionnaire. The preloaded variables represent gender, birth date, and enrollment information as it existed in the Phase I files at the time the CATI interviews were conducted. Subsequent cleaning of the Phase I data caused some of these values to change. The values for gender and birth date were verified during the Phase II interview. If the respondent corrected these values, the change was reflected in subsequent releases of the Phase I data files, as well as in the Phase II data files. The latest version of the Demographic File represents the best current information available to the study from all sources on participants' demographics. The analyst should always check the release number of a file when there are apparent discrepancies. Also note that it is AHS policy that release numbers of the underlying data files be included in all reports of study data.

Some information, such as facts about farm wells, was not checked across questionnaires at the time of the interviews. Differences in responses to such questions may represent changes of facts (a new well may have been drilled) or simply changes in responses by respondents that reflect acquisition of new knowledge, lapses in memory at the time of one of the data collection events, or any number of other reasons. Such discrepancies are to be expected in a series of lengthy questionnaires administered approximately 5 years apart. If the number and size of such differences are small, they are likely to have minimal impact on standard errors. If there are many such differences, or if the magnitude is large for a specific variable used in an analysis, the investigator should consider the potential implications for any results and whether or not it is worthwhile attempting to collect additional information from the respondents.

Enrollment dates were preloaded into CATI from the Phase I data files. There were 191 instances in which the enrollment dates were corrected on the Phase I files after the incorrect preloaded dates were used in the CATI interview. These dates were used as the basis for the following questions (Private Applicator Questionnaire item numbers are preceded by a "P," Spouse Questionnaire item numbers are preceded by an "S"):

- S4a. Was your residence in (year of enrollment) a farm? (i.e., was your house located in a farming operation that sold over \$1,000 worth of crops, nursery, greenhouse, or animal products in a year)?

- P10/S10. Have you been actively engaged in farming or worked on a farm during any of the years since (year of enrollment)?
- P11. Since (date of enrollment), did you personally mix, load, or apply pesticides for use on crops, animals, or for any other purpose?
- P21. Since (enrollment date) have you been unable to harvest or sell all or part of the crops grown or stored on your farm because of molds or fungus?
- P39/S36. Since (year of enrollment), did you have any incidents with fertilizers, herbicides or other pesticides that caused you an unusually high personal exposure?
- P39a/S36a. How many of the exposure incidents have you had since (enrollment date)?

Private Applicator items 41-53 were preceded with the statement, “For the next set of questions, I’d like you to think about the time period from (date of enrollment) to the present time.” The corresponding items in the Spouse Questionnaire were numbered 38-50. These questions are:

- P41/S38. After mixing or applying pesticides, do (did) you usually bathe or shower before continuing with other farm activities?
- P42/S39. Do (did) you usually wear the same work clothes you used to mix or apply pesticides two or more days without washing them?
- P43/S40. Do (did) you usually take your work boots off before entering your home?
- P44/S41. Are (were) the clothes you use(d) when mixing or applying pesticides usually washed separately?
- P45/S42. Are (were) agricultural or commercial pesticides ever stored (even temporarily) in your home or basement?
- P46/S43. How far is your house from the nearest field or orchard where pesticides are applied?
- P47/S44. Does (did) the farm vehicle you usually use to apply pesticides have an enclosed cab?
- P48/S45. Do (did) you usually spray with the windows closed?
- P49/S46. Does (did) your cab have a charcoal filter?
- P50/S47. Do (did) you usually repair your own spraying or mixing equipment?
- P51/S48. In most of your fields, how often do (did) you rotate the crops that you grow?
- P51a/S48a. How often is (was) that?

P52/S49. For most of the herbicides, insecticides and fungicides that you use(d), do (did) you...

1. usually apply less than the minimum recommended label rate
2. usually apply more than the recommended label rate
3. usually apply the recommended rate

P53/S50. Have you substituted biological or cultural practices (resistant varieties, pheromones, Bt sprays, tillage, planting/harvesting date adjustments, etc.) for any of your chemical pesticide applications?

The following questions relating to enrollment year were asked only in the Spouse Questionnaire:

S51. Since (year of enrollment), has your house been treated for termites?

S52. How many times has this house been treated (do not include inspections for termites)?

S53. Since (year of enrollment), has this house been treated for flies, fleas, cockroaches, ants, or insects other than termites?

S53a. How often is this house usually treated?

The use of inaccurate enrollment dates for these interviews had minimal impact on the responses. Of the 191 incorrect dates, 147 were earlier dates in CATI than they should have been. This would not have affected the information collected in response to Questions 10 and 11, which are critical for routing the interviewer through the questionnaire and for determining the reference year, which is referred to by most of the items in the questionnaire.

For the 44 respondents with later dates preloaded in CATI than the actual enrollment dates, all the dates were 1 year later than they should have been. This group may have missed some positive responses to questions of the form “Since <DATE> have you done X?” One of these 44 was a male applicator and the remaining 43 were female spouses.

5.3 Unusual Values

As with any large-scale survey, respondents sometimes gave unusually high or low values in response to questions. Many of these were screened out during the interview by soft and hard range checks in the interview instrument. When a soft range check was violated, the interviewer was informed and either corrected his or her keying or asked the respondent to confirm the response. Hard range checks blocked the entry of out of range values.

The hard range checks were repeated in SAS as part of the editing process and any discrepant values were reported to the field stations whose staff reviewed them and either corrected an error or set an override flag indicating to the edit program that the unusual value should be accepted. Similarly, outliers as defined by Tukey's (1977) method were brought to the attention of the field stations for a number of selected variables. Field station staff reviewed a sample of these, in some cases listening to tapes of the original interviews, and rarely found a need to make a change.

While we are confident that unusual values accurately represent what the respondent told the interviewer, analysts should still check for outliers in their analysis variables and consider what, if any, action they need to take to ensure that their analyses are not adversely affected. One set of covariates of interest which has a large number of outliers is the responses to the drinking questions. For instance, there are 11 private applicators who stated that they had 30 or more drinks on a typical weekend. One respondent claimed 73 drinks.

5.4 Interpretation of Missing Data Patterns

Data may have missing values because responses to initial questions caused the interview to branch around some questions, because the respondent stated that he or she did not know an answer, because the respondent refused to answer a question, because the respondent ended the interview early, or because of anomalies introduced while editing the CATI responses. All missing value codes for numeric values are coded with a SAS missing value. This value is ".D" when the respondent's response was "Don't know." The value is ".R" when the respondent refused to answer a question. All other numeric missing value codes are a period (".") with no further distinction. Missing value codes for character variables are blanks.

5.5 Topics Discussed in Phase I User Manual that May Be of Interest

There are a number of topics discussed in the *AHS Phase I Data File Users Manual* that are also pertinent to analyses of the Phase II AHS data.:

- Identifying Appropriate Reference Groups (see Phase I Manual, Section 5.3)
- Pesticide Grouping Analyses (see Phase I Manual, Section 5.4)
- Pesticides Used in Combination with Other Pesticides (see Phase I Manual, Section 5.7)

6. REFERENCES

The following publications were referenced in the text:

Crop protection handbook. (2003). Meister Pub. Co., Willoughby, OH.

Dosemeci M., Alavanja M. C. R., Rowland A. S., Mage D., Zahm S. H., Rothman N., Lubin J. H., Hoppin J. A., Sandler D. P., Blair A. (2002) A semi-quantitative approach for estimating exposure to pesticides in the Agricultural Health Study. *Annals of Occupational Hygiene*, 46(2), 245-260.

Tukey, J.W. (1977) *Exploratory Data Analysis*, Addison-Wesley, Reading, MA.

U.S. Environmental Protection Agency, Pesticide Product Information System (PPIS), web site: <http://www.epa.gov/opppmsd1/PPISdata/index.html>

The following publications contain a small selection of Agricultural Health Study papers that have been published. This list is constantly growing. For an up-to-date list of AHS publications, visit our Web page at <http://www.aghealth.org/publications.html>. As of December 2006, it contains references to more than 70 publications and has links to abstracts and full text versions of the articles for most of them.

Alavanja, M. C., Sandler, D. P., McMaster, S. B., Zahm, S. H., McDonnell, C. J., Lynch, C. F., Pennybacker, M., Rothman, N., Dosemeci, M., Bond, A. E., & Blair, A. (1996). The Agricultural Health Study. *Environ Health Perspect*, 104(4), 362-369.

Blair A, Tarone R., Sandler D., Lynch C. F., Roland A., Wintersteen W., Dosemeci, M., Alavanja M. C. R. (2002). Reliability of reporting on lifestyle and agricultural factors by a sample of participants in the Agricultural Health Study from Iowa. *Epidemiology* 13:94-99.

Hoppin J. A., Yucel F., Dosemeci M., Sandler D. P. (2002) Accuracy of self-reported pesticide use duration information from licensed pesticide applicators in the Agricultural Health Study. *J Expo Anal Environ Epidemiol*, 12:313-318.

7. QUESTIONNAIRES

Copies of the final version of each of the following questionnaires can be found on the Agricultural Health Study website at:

- Pesticide Use Modules
 - Private Pesticide Applicator
 - Spouse
 - Commercial Pesticide Applicator
- Health Modules
 - Health Module for Men
 - Follow-up Health Questionnaire for Women (Version A)
 - Follow-up Health Questionnaire for Women (Version B)
 - Follow-up Health Questionnaire for Women (Version C – Applicator)
 - Follow-up Health Questionnaire for Women (Version C – Spouse)
- Cooking Practices Module
- Diet History Questionnaire

8. CODEBOOKS

Codebooks for each data file are included in the appendices. 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:

- CATI Interview Data Files
 - Private Applicator File
 - Private Applicator Supplemental Pesticide File
 - Spouse File
 - Spouse Supplemental Pesticide File
 - Commercial Applicator File
- Diet History Questionnaire File
- Analysis Files
 - Private Applicator Base File
 - Private Applicator Pesticide Usage File
 - Private Applicator Health File

Appendix A
Pesticides Included on A List

A-LIST PESTICIDES, GROUPED BY FUNCTIONAL CLASS

Active Ingredient	PC Code	Phase II Indicator Name	Phase I Indicator Name
50 KEY VARIABLES FROM PHASE I			
Methyl Bromide	053201	A2_FUMIGANT_RY_A_1	A_FUMIGANT_CD1
Aluminum Phosphide	066501	A2_FUMIGANT_RY_A_2	A_FUMIGANT_CD2
Carbon Tetrachloride/ Carbon Disulfide	016501	A2_FUMIGANT_RY_A_3	A_FUMIGANT_CD3
Ethylene-Dibromide	042002	A2_FUMIGANT_RY_A_4	A_FUMIGANT_CD4
Benomyl	099101	A2_FUNGICIDE_RY_A_1	A_FUNGICIDE_CD1
Chlorothalonil	081901	A2_FUNGICIDE_RY_A_2	A_FUNGICIDE_CD2
Captain	081301	A2_FUNGICIDE_RY_A_3	A_FUNGICIDE_CD3
Maneb	014505	A2_FUNGICIDE_RY_A_4	A_FUNGICIDE_CD4
Metalaxyl	113501	A2_FUNGICIDE_RY_A_5	A_FUNGICIDE_CD5
Ziram	034805	A2_FUNGICIDE_RY_A_6	A_FUNGICIDE_CD6
Atrazine	080803	A2_HERBICIDE_RY_A_1	A_HERBICIDE_CD1
Dicamba	029801	A2_HERBICIDE_RY_A_2	A_HERBICIDE_CD2
Cyanazine	100101	A2_HERBICIDE_RY_A_3	A_HERBICIDE_CD3
Chlorimuron-Ethyl	128901	A2_HERBICIDE_RY_A_4	A_HERBICIDE_CD4
Metolachlor	108801	A2_HERBICIDE_RY_A_5	A_HERBICIDE_CD5
EPTC	041401	A2_HERBICIDE_RY_A_6	A_HERBICIDE_CD6
Alachlor	090501	A2_HERBICIDE_RY_A_7	A_HERBICIDE_CD7
Metribuzin	101101	A2_HERBICIDE_RY_A_8	A_HERBICIDE_CD8
Paraquat	061603	A2_HERBICIDE_RY_A_9	A_HERBICIDE_CD9
Petroleum Distillates	063501	A2_HERBICIDE_RY_A_10	A_HERBICIDE_CD10
Pendimethalin	108501	A2_HERBICIDE_RY_A_11	A_HERBICIDE_CD11
Imazethapyr	128922	A2_HERBICIDE_RY_A_12	A_HERBICIDE_CD12
Glyphosate	417300	A2_HERBICIDE_RY_A_13	A_HERBICIDE_CD13
Silvex	082501	A2_HERBICIDE_RY_A_14	A_HERBICIDE_CD14
Butylate	041405	A2_HERBICIDE_RY_A_15	A_HERBICIDE_CD15
Trifluralin	036101	A2_HERBICIDE_RY_A_16	A_HERBICIDE_CD16
2,4-D	030001	A2_HERBICIDE_RY_A_17	A_HERBICIDE_CD17
2,4,5-T	082001	A2_HERBICIDE_RY_A_18	A_HERBICIDE_CD18

Active Ingredient	PC Code	Phase II Indicator Name	Phase I Indicator Name
Permethrin	109701	A2_INSECTICIDE_RY_A_1	A_INSECTICIDE_CD1
Permethrin	109701	A2_INSECTICIDE_RY_A_2	A_INSECTICIDE_CD2
Terbufos	105001	A2_INSECTICIDE_RY_A_3	A_INSECTICIDE_CD3
Fonofos	041701	A2_INSECTICIDE_RY_A_4	A_INSECTICIDE_CD4
Trichlorfon	057901	A2_INSECTICIDE_RY_A_5	A_INSECTICIDE_CD5
Lindane	009001	A2_INSECTICIDE_RY_A_6	A_INSECTICIDE_CD6
Carbofuran	090601	A2_INSECTICIDE_RY_A_7	A_INSECTICIDE_CD7
Chlorpyrifos	059101	A2_INSECTICIDE_RY_A_8	A_INSECTICIDE_CD8
Malathion	057701	A2_INSECTICIDE_RY_A_9	A_INSECTICIDE_CD9
Parathion	057501	A2_INSECTICIDE_RY_A_10	A_INSECTICIDE_CD10
Carbaryl	056801	A2_INSECTICIDE_RY_A_11	A_INSECTICIDE_CD11
Diazinon	057801	A2_INSECTICIDE_RY_A_12	A_INSECTICIDE_CD12
Aldicarb	098301	A2_INSECTICIDE_RY_A_13	A_INSECTICIDE_CD13
Phorate	057201	A2_INSECTICIDE_RY_A_14	A_INSECTICIDE_CD14
Aldrin	045101	A2_INSECTICIDE_RY_A_15	A_INSECTICIDE_CD15
Chlordane	058201	A2_INSECTICIDE_RY_A_16	A_INSECTICIDE_CD16
Dieldrin	045001	A2_INSECTICIDE_RY_A_17	A_INSECTICIDE_CD17
DDT	029201	A2_INSECTICIDE_RY_A_18	A_INSECTICIDE_CD18
Heptachlor	044801	A2_INSECTICIDE_RY_A_19	A_INSECTICIDE_CD19
Toxaphene	080501	A2_INSECTICIDE_RY_A_20	A_INSECTICIDE_CD20
Coumaphos	036501	A2_INSECTICIDE_RY_A_21	A_INSECTICIDE_CD21
Dichlorvos	084001	A2_INSECTICIDE_RY_A_22	A_INSECTICIDE_CD22
ADDITIONAL PESTICIDES OF INTEREST			
Mancozeb	014504	A2_FUNGICIDE_RY_AA_1	NA
Acetochlor	121601	A2_HERBICIDE_RY_AA_1	NA
Acifluorfen	114401	A2_HERBICIDE_RY_AA_2	NA
Ametryn	080801	A2_HERBICIDE_RY_AA_3	NA
Bentazone	275200	A2_HERBICIDE_RY_AA_4	NA
Bromoxynil	035301	A2_HERBICIDE_RY_AA_5	NA
Clopyralid	117403	A2_HERBICIDE_RY_AA_6	NA
Cloransulam-Methyl	129116	A2_HERBICIDE_RY_AA_7	NA
S-Dimethenamid	120051	A2_HERBICIDE_RY_AA_8	NA
Fenoxaprop-P-Ethyl	129092	A2_HERBICIDE_RY_AA_9	NA
Fluazifop-Butyl	122805	A2_HERBICIDE_RY_AA_10	NA
Fluazifop-P-Butyl	122809	A2_HERBICIDE_RY_AA_11	NA
Flumetsulam	129016	A2_HERBICIDE_RY_AA_12	NA

Note: There are two entries for Permethrin. In Phase I a distinction was made between Permethrin for crops and Permethrin for poultry, livestock, or animal confinement areas. In Phase II there is no such distinction. The second permethrin line has been kept as a placeholder in the array.

Active Ingredient	PC Code	Phase II Indicator Name	Phase I Indicator Name
Glufosinate-Ammonium	128850	A2_HERBICIDE_RY_AA_13	NA
Imazaquin	128848	A2_HERBICIDE_RY_AA_14	NA
Isoxaflutole	123000	A2_HERBICIDE_RY_AA_15	NA
Linuron	035506	A2_HERBICIDE_RY_AA_16	NA
Nicosulfuron	129008	A2_HERBICIDE_RY_AA_17	NA
Sethoxydim	121001	A2_HERBICIDE_RY_AA_18	NA
Simazine	080807	A2_HERBICIDE_RY_AA_19	NA
Fomesafen	123802	A2_HERBICIDE_RY_AA_20	NA
Picloram	005101	A2_HERBICIDE_RY_AA_21	NA
Acephate	103301	A2_INSECTICIDE_RY_AA_1	NA
Cyfluthrin	128831	A2_INSECTICIDE_RY_AA_2	NA
Amitraz	106201	A2_INSECTICIDE_RY_AA_3	NA
Disulfoton	032501	A2_INSECTICIDE_RY_AA_4	NA
Endosulfan	079401	A2_INSECTICIDE_RY_AA_5	NA
Ethoprop	041101	A2_INSECTICIDE_RY_AA_6	NA
Pyrethrins	069001	A2_INSECTICIDE_RY_AA_7	NA
Pyrethrum	069000	A2_INSECTICIDE_RY_AA_8	NA
Tefluthrin	128912	A2_INSECTICIDE_RY_AA_9	NA
Aliphatic Petroleum Hydrocarbons	063503	A2_INSECTICIDE_RY_AA_10	NA
Rotenone	071003	A2_INSECTICIDE_RY_AA_11	NA

Appendix B

Pesticides Included on B List

B-LIST PESTICIDES, GROUPED BY FUNCTIONAL CLASS

(NOT A-LIST AND FREQUENCY >= 300 APPLICATORS)

Active Ingredient	PC Code	Phase II Indicator Name	Phase I Indicator Name
1,3-Dichloropropene	029001	A2_FUMIGANT_RY_B_1	NA
Chloropicrin	081501	A2_FUMIGANT_RY_B_2	NA
Copper Chloride	008001	A2_FUNGICIDE_RY_B_1	NA
Copper Sulfate	008101	A2_FUNGICIDE_RY_B_2	NA
Tebuconazole	128997	A2_FUNGICIDE_RY_B_3	NA
Diflufenzopyr	005108	A2_HERBICIDE_RY_B_1	NA
2,4-DB	030819	A2_HERBICIDE_RY_B_2	NA
Bromoxinyl Octanoate	035302	A2_HERBICIDE_RY_B_3	NA
Fluometuron	035503	A2_HERBICIDE_RY_B_4	NA
Sodium Bentazon	103901	A2_HERBICIDE_RY_B_5	NA
Clethodim	121011	A2_HERBICIDE_RY_B_6	NA
Flufenacet	121903	A2_HERBICIDE_RY_B_7	NA
Mesotrione	122900	A2_HERBICIDE_RY_B_8	NA
Flumatralin	123001	A2_HERBICIDE_RY_B_9	NA
Clomazone	125401	A2_HERBICIDE_RY_B_10	NA
Imazapyr	128821	A2_HERBICIDE_RY_B_11	NA
Thifensulfuron	128845	A2_HERBICIDE_RY_B_12	NA
Lactofen	128888	A2_HERBICIDE_RY_B_13	NA
Primisulfuron-Methyl	128973	A2_HERBICIDE_RY_B_14	NA
Rimsulfuron	129009	A2_HERBICIDE_RY_B_15	NA
Sulfentrazone	129081	A2_HERBICIDE_RY_B_16	NA
Ioxynil	353200	A2_HERBICIDE_RY_B_17	NA
Fatty Alcohols	079037	A2_HERBICIDE_RY_B_18	NA
Decanol	079038	A2_HERBICIDE_RY_B_19	NA
Maleic Hydrazide	051501	A2_HERBICIDE_RY_B_20	NA
Triclopyr	116001	A2_HERBICIDE_RY_B_21	NA
Ethephon	099801	A2_HERBICIDE_RY_B_22	NA
Bacillus Thuringiensis	006400	A2_INSECTICIDE_RY_B_1	NA
Phosphamidon	018201	A2_INSECTICIDE_RY_B_2	NA
Dimethoate	035001	A2_INSECTICIDE_RY_B_3	NA
Phosmet	059201	A2_INSECTICIDE_RY_B_4	NA

Active Ingredient	PC Code	Phase II Indicator Name	Phase I Indicator Name
Methomyl	090301	A2_INSECTICIDE_RY_B_5	NA
Esfenvalerate	109303	A2_INSECTICIDE_RY_B_6	NA
Lambda Cyhalothrin	128897	A2_INSECTICIDE_RY_B_7	NA
Tebupirimfos	129086	A2_INSECTICIDE_RY_B_8	NA
Imidacloprid	129099	A2_INSECTICIDE_RY_B_9	NA
Diphacinone	067701	A2_RODENTICIDE_RY_B_1	NA
Warfarin	086002	A2_RODENTICIDE_RY_B_2	NA
Bromadiolone	112001	A2_RODENTICIDE_RY_B_3	NA
Brodifacoum	112701	A2_RODENTICIDE_RY_B_4	NA

Appendix C

BMI Estimation Project

BMI DATA ESTIMATION PROJECT

By Ross Ulmer

This original goal of this project was to address the problem of large numbers of missing values for enrollment height and weight by imputing these values whenever possible from CATI and/or Iowa Drivers License height and weight data. There were approximately 38,500 missing values for enrollment height and 40,400 missing values for enrollment weight. Of these, close to 60% (approximately 23,200) had either CATI data or both CATI and DL data . (See Table 1)

However before any valid estimation/imputation could take place a number of data quality issues in both the enrollment and CATI height and weight data had to be resolved. The main types of data validity issues were:

- Within-Source Issues:
 - Implausibly (or impossibly) extreme values of height and weight. The original height values ranged from 1” to 95” (7’11”) for enrollment and 13” to 84” at CATI.
 - Suspicious frequency spikes in the distribution (e.g., a spike of 82 men with a recorded enrollment = 60”).
 - Instances where applicator and spouse enrollment height and weight data were somehow reversed or otherwise confounded.
 - Other
- Inconsistency Across Sources:
 - Inconsistent values of adult height across enrollment, CATI, and DL.
 - Implausibly large changes in weight from DL to enrollment or enrollment to CATI.

Handling the first type of data issues was mainly a task of data cleaning or data correction. Suspicious values were identified and then either corrected or removed from the data set. The variables containing the “corrected” values are identified in the dataset as the variables EHTCORR, EWTCORR, ... See Table 2.

Resolving inconsistencies across sources on the other hand is more complex in that it often required choosing among conflicting pieces of information. The task in these cases was to make a reasonable and defensible decision about what to do when faced with two or three contradictory - but possibly all reasonable – values. The results of this process are stored in the dataset as EHTR, EWTR, CHTR, CWTR, ... See Table 3.

Once the within source and across source data issues were resolved it was then possible to carry out the original goal of this project – imputing missing values for missing enrollment BMI. In the final analysis we were able to compute BMIs for 67,447 subjects. Of these 23,499 are based on EWTS estimated from CATI or CATI and DL. See Table 4.

TABLES

Table 1. Height in Inches and Weight in Pounds

<i>Original Ht and Wt Values</i>		N	Missing	Mean	Min	Max
Enrollment	HT (<i>ehtorig</i>)	46179	38563	67.65	1	95
	WT (<i>ewtorig</i>)	44324	40418	175.02	71	467
CATI	HT (<i>chtorig</i>)	56946	27796	68.27	13	84
	WT (<i>cwhtorig</i>)	55615	29127	183.51	1	560
Iowa DL	HT (<i>dht</i>)	42442	42300	68.45	56	83
	WT (<i>dwt</i>)	42442	42300	164.26	79	380

Table 2. Corrected Values

<i>Corrected Values</i>		N	Mean	Min	Max
Enrollment	HT (<i>ehtcorr</i>)	45747	67.74	55	86
	WT (<i>ewtcorr</i>)	44045	175.00	86	467
CATI	HT (<i>chtcorr</i>)	56847	68.28	55	84
	WT (<i>cwtcorr</i>)	55569	180.00	89	560

Table 3. Resolved Values

<i>Resolved Values</i>		N	Mean	Min	Max
Enrollment	HT	75984	68.37	55	82
	WT	67544	179.04	86	518
CATI	HT	56847	68.28	55	84
	WT	55569	180.00	89	560

Table 4. Final BMI Values

<i>Final BMI Values</i>		N	Mean	Min	Max
Enrollment	BMI (<i>ebmir</i>)	67457	26.87	13	76