Agricultural Health Study

Phase I

Data File Users Manual

Version P1REL0906.00

Stanley E. Legum, Ph.D., Editor

Michael C. R. Alavanja, Dr. P.H., NCI Project Officer
Dale P. Sandler, Ph.D., NIEHS Project Officer

Phase I Data Working Group

Susan Acker
Michael C. R. Alavanja, Dr. P.H.
Larry Engel, Ph.D.
Jane Hoppin, Sc.D.
Stanley E. Legum, Ph.D.
Stuart Long
Dale P. Sandler, Ph.D.
Marsha Shepherd

July 2009

TABLE OF CONTENTS

Chapter		Page
1	INTRODUCTION	1-1
2	DATA COLLECTION PROCEDURES	2-1
	2.1 Enrollment at Annual Pesticide Certification Sessions	2-1
	2.2 Take-Home Questionnaires	2-2
	2.3 Calls to Increase Response	2-2
	2.4 Calls to Collect Key Data Missing from Returned Questionnaires	2-3
	2.5 Validation Studies	2-4
	2.5.1 Young Women's Health Study (YWH)	2-4
	2.5.2 Neurologic and Immunologic Disease Study (NID)	2-5
	2.6 Women's Health Study (WHS)	2-5
	2.7 Rules for Handling Questionnaire Problems	2-5
3	DATA PROCESSING PROCEDURES	3-1
	3.1 Optical Scanning of Questionnaires	3-1
	3.2 Data Cleaning	3-2
	3.2.1 Initial Cleaning	3-2
	3.2.2 Final Data Cleaning	3-3
	3.3 Preparation of Analytic Files	3-3
	3.4 Data Values for Smoke Years and NSAID Variables	3-5
4	DATA FILE DESCRIPTIONS	4-1
	4.1 Private Applicator File	4-2
	4.2 Commercial Applicator File	4-2
	4.3 Spouse File	4-3
	4.4 Female and Family Health File	4-3
	4.5 Supplemental Spouse File	4-4
	4.6 Verbatim Response Files	4-5
	4.7 Validation Study Files	4-6

TABLE OF CONTENTS (continued)

Chapter				Page
5	USAGE NOTES			5-1
	5.1	How to Link Files for Analyses		5-1
	5.2	Definition of Exposure Measures		5-2
		5.2.1	Duration	5-2
		5.2.2	Frequency	5-3
		5.2.3	Cumulative Exposure	5-3
		5.2.4	Intensity	5-4
		5.2.5	Intensity-Adjusted Cumulative Exposure	5-5
	5.3	Identifying Appropriate Reference Groups		5-6
	5.4	Pesticide Grouping Analyses		5-6
	5.5	Use of Midpoint Codes		5-12
	5.6	Analysis of Fungicide Duration and Frequency Responses		5-12
	5.7	Pesticides Used in Combination with Other Pesticides		5-13
	5.8	Unusual Values		5-13
	5.9	Medical Condition Variables		5-14
	5.10	Interpretation of Missing Data Patterns		5-16
		5.10.1	Missing Value Codes	5-17
		5.10.2	Questionnaires Completed by Phone	5-17
		5.10.3	Questionnaires Collected during Reliability Study	5-18
		5.10.4	Respondent Missed a Page or Stopped	5-18
	5.11	SIC/SOC Coding		5-19
	5.12	Limitations and Uses of the Commercial Applicator File		5-19
	5.13	Order and Naming of Medical Condition Variables		5-20
	5.14	Applicator-Spouse Variable Crosswalk		5-22
	5.15	Use of Supplemental SAS Format and Attribute Statements		5-22
6	REFERENCES			6-1
	6.1	Methods		6-1
	6.2	Exposure Assessment		6-2
		6.2.1	High Pesticide Exposure Events	6-2
		6.2.2	Environmental Measures	6-2
	6.3	Health Outcomes		6-3
	6.4	Diet		6-3

TABLE OF CONTENTS (continued)

Chapter		Page
7	DATA FILE REQUESTS	7-1
8	QUESTIONNAIRES	8-1
9	CODEBOOKS	9-1
10	PHASE I ALGORITHM INTENSITY SCORES -- VERSION 2	10-1

List of Tables

Table

1-1 Composition of Cohort 1-1

2-1 Questionnaire Problem Resolution 2-6

4-1 Description of Validation Study Files 4-7

5-1 Variables Containing Total Cumulative Exposure for Chemical Groups 5-4

5-2 Algorithm I Intensity Variables 5-5

5-3 Algorithm II Intensity Variables 5-5

5-4 Questions Asking about Usage of Functional Classes of Pesticides in Each Questionnaire 5-7

5-5 Contribution of Enrollment and Take-home Questionnaires to the Number of Applicators Classified as Exposed to Each Class of Pesticides -- Counts .. 5-10

5-6 Contribution of Enrollment and Take-home Questionnaires to the Number of Applicators Classified as Exposed to Each Class of Pesticides -- Percents 5-11

5-7 Location and Number of Medical Condition Questions 5-15

5-8 Questions in the Farmer Questionnaire and Commercial Applicator Questionnaire with Wording Differences 5-20

7-1 Sample from Applicator_vars.xls 7-2

9-1 Codebook File Names 9-2

10-1 Application Methods to Be Considered for Application Score According to Pesticide Class 10-3

10-2 Application Scores Given to Each Method and Variables Indicating Use of the Method. 10-3

1. INTRODUCTION

The Agricultural Health Study (AHS) 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 AHS has four major components:

1. The main prospective cohort study of cancer and noncancer outcomes that includes:

    a. Linkage with cancer registries and vital statistics registries;

    b. Ongoing data collection (i.e., telephone interview, food frequency questionnaire, and cheek cell collection (buccal cells));

2. Cross-sectional studies that include (a) questionnaire data, (b) functional measures, (c) biomarkers, and (d) geographic information system (GIS);

3. Nested case-control studies; and

4. Exposure assessment and validation studies.

The AHS cohort includes 89,656 private pesticide applicators, spouses of private applicators, and commercial pesticide applicators recruited within Iowa and North Carolina (**Error! Reference source not found.**). Data for the main cohort study have been collected in phases at approximately five year intervals. Phase I, initial cohort recruitment, began in December 1993 and concluded in 1997. Phase II followup began in 1999 and concluded for private applicators and spouses in 2003. Phase II followup of commercial applicators started in October 2003 and concluded in October 2005. The Phase III followup began in November 2005 and is scheduled to conclude in 2010.

Table 1-1. Composition of Cohort

Type of Respondent	Number Enrolled
Private Applicators	52,394
Spouses	32,346
Commercial Applicators	4,916
Total	89,656

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, like 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).

This manual describes the Phase I data files and provides the basic information that an analyst needs to make use of the files. 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

Data were collected both in person and by mailed scannable forms. These data were supplemented by telephone calls. Each mode of data collection is described separately below.

2.1 Enrollment at Annual Pesticide Certification Sessions

Farmers and commercial pesticide applicators were identified when they sought a restricted-use pesticide license from the state Cooperative Extension Services or Departments of Agriculture. All persons in Iowa and North Carolina who wish to apply restricted-use pesticides must obtain a pesticide applicator license by undergoing training or testing in the safe handling of pesticides. There are two licensing categories: "private" applicators (i.e., farmers) and "commercial" applicators (persons employed by pest control companies or by businesses that use pesticides but whose primary function is not pesticide application, e.g., grain millers, warehouse operators).

At the licensing facility each pesticide applicator was asked to complete a brief, Enrollment Questionnaire. In Iowa, both commercial and farmer applicators attended some of the same sessions and were invited to participate in the study. In North Carolina, farmers and commercial applicators attended separate training sessions; only farmer applicators from North Carolina were enrolled. Iowa and North Carolina Field Station staff administered and collected the questionnaires.

Approximately 300 people enrolled in IA in Year 3 using an abbreviated version of the Enrollment Questionnaire known as the "Followup Questionnaire." They picked this up instead of the Enrollment Questionnaire. Although it was intended that the respondents to this questionnaire be people who had completed an Enrollment Questionnaire one year previously, 344 new respondents completed the Followup Questionnaire but not the Enrollment Questionnaire. As a result, these individuals do not have complete enrollment information. Details of this are discussed in Section 5.10.3.

2.2 Take-Home Questionnaires

The participating applicators were also given a packet of additional questionnaires to complete at home and mail back. Farmer applicators were given three take-home questionnaires: (1) a Farmer Applicator Questionnaire, (2) a Spouse Questionnaire, and (3) a Female and Family Health Questionnaire to be completed by the wife of a male farmer or by the farmer, if female. Commercial applicators were given the Commercial Applicator Questionnaire[1] to complete and mail. It was nearly identical to the Farmer Applicator Questionnaire except for modifications that removed questions about farming practices and a question about the distance of the farm's well from fields where pesticides were applied. Female commercial applicators were also given a Female and Family Health Questionnaire to complete at home.

Copies of all the questionnaires are located at the end of this manual.

2.3 Calls to Increase Response

To boost enrollment, both Field Stations conducted telephone interviews of nonenrolled spouses of enrolled private applicators using the scannable Spouse Questionnaire from the take-home packet. For telephone administration, Sections I-V of the Spouse Questionnaire were administered first, followed by Sections IX and X. After Section X, the personal identifiers were requested. If there was time left, or the respondent was willing to go beyond the initial limit of 30 minutes, Sections VI-VIII were administered. Some respondents elected to complete the full questionnaire through a mail administration. Some spouses also consented to complete the Female and Family Health Questionnaire via mail. When available, completed mail questionnaires replaced data collected during telephone interviews.

Some respondents completed the take-home questionnaire but either left their Enrollment Questionnaires blank or did not turn them in. A total of 3,917 respondents completed the Enrollment Questionnaire by phone. Of these 3,902 were private applicators and 15 were commercial applicators. Those completing the questionnaire by phone are identified by the variable APPHONE. The 6,232 spouses who completed the Spouse Questionnaire by telephone are identified by the variable SPPHONE. In addition, 12 spouses who were also applicators completed the Spouse Questionnaire by phone. Their responses to the Spouse Questionnaire are

[1] The Farmer Applicator Questionnaire and the Commercial Applicator Questionnaire may also be referred to as the Farmer Questionnaire and the Commercial Questionnaire in the text.

contained in the Supplemental Spouse File (see Section 4.5). These records are also identified by the variable SPPHONE.

The telephone-administered questionnaires were recorded on the same scannable forms as were used by other respondents and were processed the same way as those completed in person or through the mail. Once the data were received at the coordinating center they were subjected to the same data and logic checks as other questionnaire data.

2.4 Calls to Collect Key Data Missing from Returned Questionnaires

During the administration of the Enrollment Questionnaire, the applicator was often rushed or did not fully understand the importance of his or her participation in the AHS. Thus, many of the Enrollment Questionnaires were not completely filled in. To have the most complete complement of data available for analysis, the most important questions for analysis on the Enrollment Questionnaire were identified. Short, customized questionnaires were developed that included only the questions that were missing from each respondent's Enrollment Questionnaire. Missing Data Questionnaires were also designed and administered to spouses of enrolled applicators.

The full Enrollment Missing Data Questionnaire consisted of 21 key questions from the Enrollment Questionnaire. One of these asked about usage of six pesticides of particular interest: 2,4-D; glyphosate products; imazethapyr products; atrazine products; chlorpyrifos products; and terbufos products. An applicator was asked only about pesticides on this list that he or she had skipped on the Enrollment Questionnaire. A North Carolina applicator was selected for calling if he or she had completed part of the Enrollment Questionnaire but had skipped four or more questions on the Enrollment Missing Data Questionnaire. A tailored questionnaire was completed for such an applicator that contained the subset of these questions which he or she failed to answer. Iowa applicators were selected in the same manner, except that the criterion was lowered to two skipped questions. This kept the questionnaires short enough to be usually administered within 15 minutes.

The full Spouse Missing Data Questionnaire consisted of 38 key questions from the Spouse Questionnaire. One of these questions asked about 21 distinct medical conditions. During the telephone interview, a spouse was asked about a medical condition only if she or he had skipped it on the mailed Spouse Questionnaire. North Carolina spouses who had partially completed questionnaires, but had skipped five or more of the key questions on the Spouse

Questionnaire were selected for telephone interviews. In Iowa, the threshold was two or more missing key questions from the Spouse Questionnaires. The Field Stations administered the questionnaires over the telephone.

The Coordinating Center key entered and verified the completed missing data questionnaires. These data were then merged onto the master data file corresponding to the Enrollment or Spouse Questionnaire, as appropriate.

2.5 Validation Studies

Three validation studies were conducted by the project to determine whether people not returning the take-home questionnaires demonstrated different characteristics to key questions than cohort members who mailed in the questionnaires as expected. The three studies were the Young Women's Health Study (YWH), the Women's Health Study (WHS), and the Neurologic and Immunologic Disease Study (NID).

Three separate random samples of 1000 persons were selected from the existing cohort by the Coordinating Center in Year 2 of the study. These samples, selected from among persons (or spouses of such persons) who complete the Enrollment Questionnaire, were followed by the Field Stations through usual means to obtain completed questionnaires. Those who ultimately did not comply were then contacted by the Field Stations for a brief focused telephone interview covering selected questions from either the Farmer-Applicator or the Spouse and Female and Family Health questionnaires. The validation questionnaires were double key-entered and verified at the Coordinating Center.

2.5.1 Young Women's Health Study (YWH)

Since the age of spouses who had not replied was unknown, all male private applicators who enrolled in Year 2 of data collection were identified. A subset of male applicators between the ages of 30 and 45 was created. A random sample of 1,000 applicators was selected from this subsample. (An over-sample of 20 additional applicators was selected to be used as replacements in case of duplicates, or other ineligibility factors). The spouses of 652 of these applicators were identified as nonresponders to the Female and Family Health Questionnaire. Of these, 471 were interviewed by phone using the Young Women's Health Questionnaire.

2.5.2 Neurologic and Immunologic Disease Study (NID)

All male private applicators who enrolled during Year 2 of data collection were identified. A subset of male applicators who completed an Enrollment Questionnaire and were between the ages of 40 and 69 was identified. A random sample of 1,000 applicators was selected and those who had not returned the Private Applicator Questionnaire (the take-home questionnaire for private applicators) were identified. An attempt was made to contact all of the nonresponders from Iowa but, because of budget restrictions, only 60 percent of the nonresponders from North Carolina were selected for interviewing. This 60 percent was randomly selected from the nonresponding group in North Carolina. The Field Stations attempted to interview by phone a total of 470 men who had not returned the take-home questionnaire and were able to complete 326 interviews using the Neurologic and Immunologic Disease Questionnaire.

The variables for the nonfarm job and the industry of that job were coded using the Standard Occupation Classifications (SOC), and Standard Industrial Classifications (SIC) (Major Titles). The SIC/SOC code and definition were added as two additional fields at the end of the data.

2.6 Women's Health Study (WHS)

As for the Young Women's Health Study, because the age of spouses who had not replied was unknown, a subset was created of all male applicators between the ages of 40 and 69 who enrolled in Year 2 of data collection. A random sample of 1,000 applicators was selected from this subsample. (An over-sample of 20 additional applicator was selected to be used as replacements in case of duplicates, or other ineligibility factors). The spouses of 558 of these applicators were identified as nonresponders to the Female and Family Health Questionnaire. Of these, 350 were interviewed by phone using the Women's Health Questionnaire.

2.7 Rules for Handling Questionnaire Problems

A number of problems arose during the data collection relating to respondents completing the wrong questionnaires or completing multiple questionnaires. A set of data

resolution rules was developed to systematize the way in which these problems were handled. In Table 2-1, the questionnaires are referred to using the following abbreviations:

Q0	Enrollment Questionnaire (completed at the licensing site)
Q1	Farmer Applicator Questionnaire or Commercial Applicator Questionnaire (take-home)
Q1A	Farmer Questionnaire (take-home)
Q1ANF	Commercial Questionnaire (take-home)
Q1B	Spouse Questionnaire
Q1H	Female and Family Health Questionnaire

Table 2-1. Questionnaire Problem Resolution

PROBLEM	RESOLUTION
No match between Q36 (on Q0) and packet type.	Called, identified real applicator type if different from packet received. If respondent had not started the packet, sent respondent correct packet; if respondent had started, encouraged him to complete it (we will accept the mismatch).
Respondent completed both commercial and private.	Respondent accepted as private and commercial Q1ANF data transferred to Farmer Applicator Questionnaire. Q1B and Q1H mailed to married applicators for completion by spouse.
Switched packets with both applicators known	Wrote new numbers on both Q0s, initial, and document.
Unmarried applicator completed Q1B but not Q1A.	Q1B scanned in normal manner. Asked applicator to complete Q1A.
Applicator completed Q1H, but spouse did not.	Asked spouse to complete Q1H. When spouse completed Q1H, deleted the Q1H the applicator filled in from database. If spouse declined participation, left Q1H completed by applicator in database and sent to scanning.

Table 2-1. Questionnaire Problem Resolution (continued)

PROBLEM	RESOLUTION
Applicator refused Q0, but completed a Q1A and spouse completed Q1B and/or Q1H.	Q1 questionnaires kept in data. A blank Q0 will be used in the enrollment database for ID linkage to the Q1s.
Duplicate Q0s	The bar code ID of the earliest enrollment was used even if it was not the most completed Q0. On a case-by-case basis, demographic and background information from the duplicate Q0s could be merged into the accepted Q0 to make a more complete document.
Duplicate Q1s	Questionnaires were evaluated for completeness. If equal, the one with the earlier enrollment date was accepted and assigned the matching bar code number of the accepted Q0. If forms were not equal, the form with the most complete information was kept and assigned the matching bar code number of the accepted Q0. The duplicate data were deleted from the database and archived.
Spouse pairs (both husband and wife were applicators)	If two Q1Bs and Q1Hs, only those completed by the female were accepted; those completed by the male were deleted and archived. If, the Q1B completed by the male was definitively more complete than the Q1B completed by the female, the Q1B completed by the male was accepted. The Q1B was placed under the male ID and the Q1H was placed under the female ID. If necessary, flags were placed in the enrollment data to indicate where the Q1B or Q1H can be located.
Switched packet (respondent completed Q1 packet with different ID than his/her enrollment form)	Field Station identified the changes and notified Coordinating Center of ID changes. These changes were made to enrollment database ONLY.
Wrong respondent (Q1A completed by person different from enrollment form)	Respondent completing the Q0 was contacted to complete the Q1A. If respondent did complete the Q1A, the previous Q1A was replaced with the new Q1A (REDO). If the applicator refused to complete the Q1A, the information completed by the other person was sent to scanning but was noted as “wrong respondent” at the Field Station.

3. DATA PROCESSING PROCEDURES

There were three major stages in processing the data:

- Optical scanning of questionnaires,
- Data cleaning, and
- Preparation of analytic files.

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

3.1 Optical Scanning of Questionnaires

The questionnaires were optically scanned. Quality control checks were performed before and after each scanning run to ensure that the scanning equipment was operating correctly and that all pages put through the scanner were actually scanned. A computer program identified items in a questionnaire that contained invalid values or multiple responses. All such responses were reviewed by a data editor and corrected if possible. If an editor could not determine what an item should be, it was left as scanned. For example, if an item was designed to allow only one response but two were marked, an editor would examine the questionnaire. If one of the marks had been clearly erased, the editor made the appropriate change. If on the other hand, the two marks were similar, the editor left the response coded as a multiple response.

After the questionnaires had been scanned and edited, any text which was to be key entered was keyed. The scanning program detected responses that contained handwriting and brought these items to the attention of the key entry staff who then entered them from the source document into the scanning data file.

Scanning was performed in batches of questionnaires of the same type on a flow basis.

3.2 Data Cleaning

Data cleaning was performed in order to produce a set of clean data files that reflected as accurately as possible what respondents intended to communicate while completing the questionnaires.

3.2.1 Initial Cleaning

An extensive set of edit checks was developed to test for unusual patterns among the responses to a questionnaire. These checks were incorporated into a computer program that was run on each batch of scanned data before the data were added to the cumulative data file. When certain edit failures were encountered, the program corrected the problem and set a flag indicating that the record had been adjusted. For instance, if a participant indicated that he did not use pesticides but later indicated that he used the herbicide Roundup, then the variable indicating that he used pesticides was changed from "No" to "Yes." Similarly, if a respondent marked "No" to indicate that he did not use 2,4-D but also indicated that he personally applied 2,4-D 5 9 days per year for 2 to 5 years, the "No" response was changed to "Yes." Other types of edit failures were listed in a report so that the original questionnaire could be examined to see if they could be corrected. When a correction could be determined from a review of the questionnaire, the revised data were input and the correction was noted on the questionnaire and in a hard-copy log.

Since birth and enrollment dates were key in determining age at enrollment, special efforts were made to ensure that these two dates were as complete and accurate as possible. A number of age-related questions in the questionnaires could be compared against a person's calculated age from the "Today's Date" question and the birth date question to identify questionnaires possibly needing correction for one of these dates.

Some respondents filled in erroneous years for the questionnaire completion date. This was particularly prevalent in January and February, and easy for the Field Stations to correct using their data collection logs. The Field Stations determined suspect or missing birth date information by utilizing such outside sources of data as driver license records and certification databases. Birth dates were also verified and corrected as needed during Phase II telephone interviews. The enrollment dates and birth dates of pesticide applicators and their spouses in these files represent the best information available to the Field Stations as of April 2006.

3.2.2 Final Data Cleaning

After the cumulative data had been compiled and the special cleaning operations were completed, a check of the files showed that a number of the initial edits were still not being passed by all the records. An extensive review was made of the edit rules, with particular attention applied to the known failure patterns. A sample of original documents was reviewed to check for residual scanning errors. No scanning errors were found during this review. Documents with multiple responses to a question were also reviewed in an attempt to determine what the respondent had intended to reply. When a clear determination could be made, an update record was created in order to correct the questionnaire data file.

A committee of data processing professionals and Agricultural Health Study analysts then reviewed and revised the edit rules to automatically correct as many of the remaining anomalies as could be resolved in an algorithmic manner.

Less than 5 percent of the records and less than 5 percent of the questions needed to be modified after the data were originally scanned. Well under 1 percent of the data needed changing during the final cleaning operations.

3.3 Preparation of Analytic Files

Analysis files were prepared from the clean data files by organizing the data in a manner that would be more easily used by analysts. The following major changes were made to the files:

- Derived variables were created and added to the files. These included estimates of intensity of exposure for pesticide classes and intensity adjusted cumulative exposure measures for each pesticide. They also included SIC and SOC codes for nonfarm occupations reported by farmers and their spouses. (SIC and SOC coding was not performed for commercial applicators who reported nonpesticide application jobs.) In addition a number of derived variables were created and added to the files. For example, some of the variables added to the Applicator File are the calculated age at enrollment; the number of alcoholic drinks consumed per month; body mass index; flags indicating whether each of a broadly defined set of pesticides was ever used (e.g., organochlorines); and estimated levels of exposure to each pesticide specified based on the number of days that an applicator applied or mixed the pesticide, protective equipment used, and personal hygiene procedures used. The full set of derived variables is documented in detail in the code book for each data file.

- Data from the Enrollment and Farmer Applicator Questionnaires were combined into a single file called the Private Applicator File. This was done in large part so that all the pesticide use questions for an applicator would appear in one file. This structure had the added benefit that the pesticide questions could now be named in a parallel manner to those in the Spouse File.

- Data from the Enrollment and Commercial Applicator Questionnaires were combined into a single file called the Commercial Applicator File. These data were combined for the same reasons as the Enrollment and Farmer Questionnaire data were combined. The Private Applicator File and the Commercial Applicator File have the same structure and can be concatenated if desired. They are kept separate because there are a number of subtle differences between the files that the analyst should keep in mind if they are to be analyzed together. These differences include the fact that there are some questions (e.g., the distance of the household well from where pesticides are applied) that were asked only in the farmer questionnaire and the fact that no spouses of commercial applicators have been enrolled in the study. Analysts should also note that SIC and SOC coding has been completed on responses to work done off the farm for farmers, but not for work other than pesticide application for the commercial applicators.

- Variables in the Applicator Files were renamed and reordered to obtain parallelism between Applicator and Spouse Files. This parallelism will simplify the programming in certain types of analyses. Variables in the Spouse File were in some instances also renamed in order to make the naming conventions more consistent.

- The values for many categorical variables were recoded to make them easier for analysts to use. No meanings were changed as a result of this recoding. For instance, gender codes were changed to 1=Male and 2=Female. Yes/No codes were changed so that 0=No and 1=Yes. In questions that asked the respondent to "mark all that apply," the codes were set so that 0=Unmarked and 1=Marked.

- Missing values were, in general, set to the standard SAS missing value code of ".". The special missing value code of ".N" was assigned to responses of "None" to the question "In an average year when you personally used this pesticide, how many days did you [apply] it?" The "None" response was available only for fungicide-related questions on the Enrollment, Farmer Applicator, and Commercial Applicator Questionnaires. The ".N" code was used only if the respondent indicated that he or she did use a fungicide, but then indicated that it was for "None" days per year. The missing value code of ".M" was used for multiple responses and the code ".U" was used for invalid responses to the schooling variable in the Spouse File (see Section 5.10.1).

3.4 Data Values for Smoke Years and NSAID Variables

In order to make the files easier to use in analyses, the following global changes were made to the datasets:

- For current or former smokers with 0 smoke years, smoke years (ASMOKYRS in the Applicator File and SSMOKYRS in the Spouse File) has been changed from 0 to 0.5.

- Changes to NSAID variables: The values of AMEDIC1 – AMEDIC7 have been updated from missing or 0 to 1 when respondent is currently taking the drugs or number of years taken is more than 1. The values of variables AMEDYR1 – AMEDYR7 have been updated from missing to 0 when AMEDIC1 – AMEDIC7 is 0. Values of variables AMEDNOW1 – AMEDNOW7 have been set to 2 (never used) when the value of AMEDIC1 – AMEDIC7 for the corresponding drug is 0.

4. DATA FILE DESCRIPTIONS

The Agricultural Health Study Phase I data files contain the content of the questionnaires completed by the participants and some derived variables based on their responses. Identifying data such as names, Social Security numbers, addresses, and phone numbers are maintained at the Field Stations which collected the data, but are not included in these data files. There are seven main data files:

- Private Applicator File;
- Commercial Applicator File;
- Spouse File; and
- Female and Family Health File;

In addition, there are a number of supplementary data files:

- Supplemental Spouse File;
- Verbatim Response Files; and
- Validation Study Files.

Each of these files is described in a separate section below.

There are also four data files containing data collected over multiple phases that analysts may want to consider when analyzing AHS data:

- Demographic Data File.
- Cancer Registry File;
- Mortality Data File; and
- Buccal Receipt File.

These four files are described in the Demographic, Registry, and Biospecimen Files User Manual.

4.1 Private Applicator File

The Private Applicator File contains the responses to the Enrollment and Farmer Questionnaires, as well as a number of derived variables (e.g., body mass index) derived from these responses. Since the study was designed for applicators to complete the Enrollment Questionnaire at their pesticide certification site and then to take home and complete either the Farmer Questionnaire (for private applicators) or the Commercial Questionnaire, no applicator is represented in both the Private Applicator File and the Commercial Applicator File (see Section 4.2). A set of flags on the file indicates precisely which questionnaires the applicator completed.

The file contains the following types of information:

- Demographic information;
- Detailed information about pesticide use;
- Lifestyle activities;
- Health information;
- General farming information;
- Work practices that might affect levels of pesticide exposure;
- Nonfarm or noncommercial pesticide applicator jobs held; and
- Dietary and cooking practices.

There are 52,394 records in the file. Each record represents one applicator.

4.2 Commercial Applicator File

The Commercial Applicator File contains the responses to the Enrollment and Commercial Questionnaires as well as the same set of derived variables as the Private Applicator File. Since the study was designed for applicators to complete the Enrollment Questionnaire at their pesticide certification site and then to take home and complete either the Farmer Questionnaire (for private applicators) or the Commercial Questionnaire, no applicator is represented in both the Commercial Applicator File and the Private Applicator File (see Section 4.1). A set of flags on the file indicates precisely which questionnaires the applicator completed.

The Commercial Applicator File has the same structure as the Private Applicator File and contains 4,916 records. Each record represents one applicator.

4.3 Spouse File

The Spouse File contains responses to the Spouse Questionnaire. This file contains data from spouses of individuals in the Applicator File. If the spouse of an applicator is a certified pesticide applicator in his or her own right, both the husband and wife are represented in the Applicator File and neither appears in the Spouse File. The reason for this is that the pesticide exposure information collected in the Enrollment Questionnaire and Farmer Questionnaire is more complete than the exposure information collected in the Spouse Questionnaire. Husband and wife pairs are explicitly identified in the Applicator File.

The Spouse File contains the following types of information.

- Demographic information;
- Information about pesticide use;
- Home and work practices;
- Lifestyle activities;
- Information about the home environment;
- Dietary and cooking practices;
- Health information; and
- Nonfarm jobs held.

There are 32,347 records in the file. Each record represents one spouse.

4.4 Female and Family Health File

The Female and Family Health File contains responses to the Female and Family Health Questionnaire. This questionnaire was completed by the female member of the farmer-spouse pair. It contains information on the woman's reproductive history, pregnancies, and children.

There are 20,249 records in the file. Each record represents one spouse.

4.5 Supplemental Spouse File

The Supplemental Spouse File contains responses to the Spouse Questionnaire which were completed in error. There were 398 married pairs of applicators. Each of these 796 respondents completed an Enrollment Questionnaire and is represented in the Private Applicator File. Of these respondents, 320 also completed the take-home Farmer Questionnaire. These responses are incorporated into the appropriate variables in the Private Applicator File. Of these 320, 97 women and 9 men also completed the Spouse Questionnaire. In order to avoid double counting of enrollees, the responses on these 106 Spouse Questionnaires are ***not*** included in the Spouse File. Since there are some unique questions in the Spouse Questionnaire that are not in the Farmer Questionnaire, we have included these records in the Supplemental Spouse File. The responses to these questions can be used by analysts as long as they are careful not to count people twice. The questions of interest are:

- Q6 – How many years have you lived on a farm? This is similar to Q2 in the Farmer Questionnaire but is numeric and provides more information than the 5- and 10-year response ranges in the Farmer Questionnaire
- Q7 – Animal contact rates
- Q16 – Number of days working in fields in last growing season
- Q18-22 – Home and work practices
- Q49-52 -- Hair coloring questions
- Q53-57 – Residence and well information
- Q58 - Q78 – House information and residential pesticide use
- Q95 - Q101 – Home grown food

There were also 95 women and 6 men who completed the Spouse Questionnaire but not the Farmer Questionnaire. Since these 101 respondents completed the Enrollment Questionnaire, they are represented in the Private Applicator File. Including them in the Spouse File would lead to double counting of the study participants. For this reason, these respondents' answers to the Spouse Questionnaire have been included in the Supplemental Spouse File instead of the regular Spouse File. Analysts can safely use these data as long as they are careful to avoid counting people twice.

The flag SUPPLEMENTAL can be used to distinguish the two types of records in the Supplemental Spouse File. A value of "1" indicates that the applicator completed both the take-home Farmer Questionnaire and the Spouse Questionnaire. A value of "2" indicates that the applicator completed the Spouse Questionnaire, but not the Farmer Applicator Questionnaire.

There are 207 records in the file. Each record represents one spouse. Since only one pair of married spouses both completed the Spouse Questionnaire, there are 206 unique participant IDs in the file. To uniquely distinguish the records in this file, it is necessary to use both the participant ID variable (PARTID) and the gender variable (SGENDER).

4.6 Verbatim Response Files

A number of items in the questionnaires required the respondents to write out their answers rather than select them from a list. These included such questions as those that asked about nonfarm employment and lists that included "Other" as a choice and then provided a place to specify the referent of "other. " We refer to responses such as these as "verbatim responses."

Verbatim responses for selected questions were flagged during the scanning of the questionnaires and were keyed. The keyed verbatim responses for each questionnaire were collected in a separate file. There are verbatim response files corresponding to each of the following questionnaires:

- Private Applicator Questionnaire (22,916 records)
- Commercial Applicator Questionnaire (2,375 records)
- Spouse Questionnaire (32,346 records)
- Supplemental Spouse Questionnaire (207 records)
- Female and Family Health Questionnaire. (20,249 records)

Note that there is no verbatim response file corresponding to the Enrollment Questionnaire. Check the codebook for each verbatim response file to see the variables that are included in the file.

4.7 Validation Study Files

The validation study interviews were each conducted as telephone interviews consisting largely of questions copied from the paper-and-pencil take-home questionnaires distributed at enrollment. In order to make the data as compatible with what is in the main questionnaire files for the Agricultural Health Study, we have made a number of changes from the original scanned data files. The major changes are:

- When the questions are identical in the validation study questionnaire and the corresponding take-home questionnaire, the variables in the validation study files have been given the same names as in the take-home questionnaires.

- The responses to the validation study questionnaires were recoded to be consistent with the response codes used in the files containing the take-home questionnaire data.

- When there is sufficient information in the validation study questionnaire to create a derived variable with the same meaning as that in the main study files, that derived variable has been added to the corresponding validation study file.

- Items in the validation study questionnaires for the Young Women's Health Study and the Women's Health Study that corresponded to the items in the Female and Family Health Study were placed in a separate file in order to simplify combining data from the Female and Family Health File with the validation study files. These files are named the YWH_FFH file and the WHS_FFH file. The remaining questions from these two validation study files are contained in the YWH_SPO file and the WHS_SPO file. These files correspond to the Spouse File, but also contain a number of additional items that are either unique to the validation studies or are similar to Spouse Questionnaire items but too different to treat as the same.

- In the NID file, if the applicator indicated that he had not used a pesticide, we set the number of years that the pesticide was used to zero. We also set the number of days per year that the pesticide was used to zero.

Table 4-1 contains a short description of each of the five validation study files and shows the number of records in each.

Table 4-1. Description of validation study files

File name	Description	Number of records
NID	The subset of questions from the Private Applicator Questionnaire that were used in the Neurologic and Immunologic Disease Questionnaire	326
YWH_FFH	The subset of questions from the Female and Family Health Questionnaire that were used in the Young Women's Health Questionnaire	350
YWH_SPO	The subset of questions from the Spouse Questionnaire that were used in the Young Women's Health Questionnaire and any questions unique to the Young Women's Health Questionnaire.	350
WHS_FFH	The subset of questions from the Female and Family Health Questionnaire that were used in the Women's Health Study Questionnaire	471
WHS_SPO	The subset of questions from the Spouse Questionnaire that were used in the Women's Health Study Questionnaire and any questions unique to the Women's Health Study Questionnaire	471

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;
- Definition of exposure measures;
- Identifying appropriate reference groups;
- Pesticide grouping analyses;
- Use of mid-point codes;
- Analysis of fungicide duration and frequency responses;
- Pesticides used in combination with other pesticides;
- Unusual values;
- Medical condition variables;
- Interpretation of missing data patterns;
- SIC/SOC coding;
- Limitations and uses of the Commercial Applicator File;
- Order and naming of medical condition variables;
- Applicator-spouse variable crosswalk; and
- Use of supplemental SAS format and attribute statements.

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

5.1 How to Link Files for Analyses

The Phase I data files are linked to each other and to Phase II data by the six-character participant ID. The participant ID can be thought of as a family ID that is used for both the applicator and his or her spouse. The ID was preprinted on each of the questionnaires

completed by the participant and his or her family. The variable AP_SPOUSE indicates whether the record is for an applicator or a spouse.

When an applicator's spouse appears in the Applicator File rather than in the Spouse File because both of them completed Enrollment Questionnaires, they have different participant IDs. Married pairs of applicators are identified and linked in the Applicator File by the SPSPAIR variable. If the SPSPAIR variable is blank, the applicator's spouse did **not** complete an Enrollment Questionnaire. If the variable contains a value, it is the participant ID assigned to the applicator's spouse on the spouse's copy of the Enrollment Questionnaire.

5.2 Definition of Exposure Measures

The questionnaires provide a wealth of self-report information that can be used to estimate levels of exposures to pesticides. The files contain five general types of exposure information: duration, frequency, cumulative exposure, intensity and intensity-adjusted cumulative exposure. Each of these measures is described in a separate section below.

5.2.1 Duration

Duration of pesticide exposure is defined as the number of years exposed to a pesticide. The applicator files (both the Private Applicator File and the Commercial Applicator File) and the Spouse File contain the response to the general question "How many years did you personally mix or apply pesticides" (Question 10a in the Enrollment Questionnaire and Question 8a in the Spouse Questionnaire). These responses are represented by variables AYRSMIX and SYRSMIX, respectively.

The Enrollment, Farmer, and Commercial Questionnaires ask the same question separately for each of 50 pesticides – 18 herbicides, 22 insecticides, 6 fungicides, and 4 fumigants.

5.2.2 Frequency

Frequency of pesticide exposure is defined as the average number of days per year that a pesticide is used. The applicator files and the Spouse File contain the response to the

general question "During those years, how many days per year did you personally apply pesticides?" (question 10b in the Enrollment Questionnaire and question 8b in the Spouse Questionnaire). These responses are represented by variables AMIXDPY and SMIXDPY, respectively.

The Enrollment, Farmer, and Commercial Questionnaires ask the same question separately for each of the listed pesticides. For example, the variable A_HERBICIDE_DAY1 contains the typical number of days per year that a respondent personally used atrazine, while the variable A_HERBICIDE_DAY2 contains similar information about dicamba usage.

5.2.3 Cumulative Exposure

Cumulative exposure is the product of duration of exposure and frequency of exposure. The cumulative exposure for any pesticide in the applicator files can be calculated by multiplying the number of years of use by the average number of days per year. For example, the cumulative exposure to chlorpyrifos is given by:

A_INSECTICIDE_YR8 x A_INSECTICIDE_DAY8

We have calculated the total cumulative exposure for each of the groups of pesticides described in Section 5.4 by summing the cumulative exposures for each of the chemicals in the group. Table 5-1 gives the names of these cumulative exposure variables for each of the chemical groups.

Table 5-1. Variables containing total cumulative exposure for chemical groups

Pesticide Group	Variable Containing Total Cumulative Days Mixing or Applying Chemicals in the Group
Carbamates	A_TOT_CUMDAYS_CAR
Fungicides	A_TOT_CUMDAYS_FNG
Fumigants	A_TOT_CUMDAYS_FUM
Herbicides	A_TOT_CUMDAYS_HRB
Insecticides	A_TOT_CUMDAYS_INS
Organochlorines	A_TOT_CUMDAYS_OCH
Organophosphates	A_TOT_CUMDAYS_OPH
Organothiophosphates	A_TOT_CUMDAYS_OTH
Phenoxys	A_TOT_CUMDAYS_PNX
Triazines	A_TOT_CUMDAYS_TRZ

5.2.4 Intensity

The intensity level of pesticide exposure is a function of pesticide handling procedures (i.e., mixing, loading, application, repair of equipment), as well as protective equipment used, and hygiene practices. Two algorithms have been defined for calculating intensity of exposure. The first is based on information collected in the Enrollment Questionnaire and is documented in Dosemeci et al. (2002). The factors used in this algorithm to adjust for different application methods and different personal protective equipment (PPE) were based on a review of the pesticide exposure. In 2002, the Environmental Protection Agency observed a subset of AHS participants applying pesticides and collected numerous biological and environmental specimens (Thomas et al, personal communication). The algorithm for calculating the intensity adjustments was revised in 2007 revised based on the results of the EPA study. The algorithm was revised again by the AHS Exposure Assessment Working Group in 2009 to reflect knowledge gained from added experience using the exposure estimates. The intensity adjusted exposures and the factors used to calculate them have similar names to the original intensity-related variables. All of the new variables have "_V2" as part of the variable name. See Chapter 10 for a detailed specification of the revised algorithm.

There are 5 intensity variables, each of which relates to a single class of pesticides that tend to be applied in a similar manner rather. These version of these variables used in the original algorithm are listed in Table 5-2. Those used in the revised version of the algorithm are listed in Table 5-3.

Table 5-2. Intensity Variables Used in Original Algorithm

Variable	Description
INT_HERB_ALG1	Herbicide Intensity (Algorithm I)
INT_CROP_ALG1	Crop Insecticide Intensity (Algorithm I)
INT_ANIM_ALG1	Animal Insecticide Intensity (Algorithm I)
INT_FUNG_ALG1	Fungicide Intensity (Algorithm I)
INT_FUMG_ALG1	Fumigant Intensity (Algorithm I)

Table 5-3. Intensity Variables Used in Revised Algorithm

Variable	Description
INT_HERB_ALG1_V2	Herbicide Intensity Score using version 2 of the algorithm
INT_CROP_ALG1_V2	Crop Insecticide Intensity Score using version 2 of the algorithm
INT_ANIM_ALG1_V2	Animal Insecticide Intensity Score using version 2 of the algorithm
INT_FUNG_ALG1_V2	Fungicide Intensity Score using version 2 of the algorithm
INT_FUMG_ALG1_V2	Fumigant Intensity Score using version 2 of the algorithm

5.2.5 Intensity-Adjusted Cumulative Exposure

The **intensity-adjusted cumulative exposure** level for a pesticide for an applicator is the product of the applicator's intensity level, duration of exposure to the pesticide, and frequency of application of the pesticide (Dosemeci et al. (2002)). Because there is a many-to-many relation between the intensity measures and the pesticides, the Applicator Files contain two derived intensity-adjusted cumulative exposure variables for each pesticide. There is one set corresponding to each of the sets of intensity level variables. Each intensity-adjusted cumulative exposure variable name has the structure:

1. A letter indicating whether it is an applicator (A) or spouse (S) variable followed by an underscore;

2. A string indicating the type of pesticide followed by an underscore: HERBICIDE_, INSECTICIDE_, FUNGICIDE_, or FUMIGANT_;

3. The characters "CUMEXP_";

4. The string "ALG1_" indicating that the exposure algorithm was used to calculate the variable;

5. For the revised version of the algorithm, the characters "V2_"; and

6. A sequence number. The pesticides have been numbered so that they are arranged in the order that they were presented in the Spouse Questionnaire.

For example the original cumulative exposure variable for Atrazine is named A_HERBICIDE_CUMEXP_ALG1_1 and the revised cumulative exposure variable is named A_HERBICIDE_CUMEXP_ALG1_V2_1. Similarly, the corresponding scores for Dicamba are A_HERBICIDE_CUMEXP_ALG1_2 (original) and A_HERBICIDE_CUMEXP_ALG1_V2_2 (revised).

5.3 Identifying Appropriate Reference Groups

Respondents were given the opportunity to mark all the responses that apply for questions regarding pesticide application methods, personal protective clothing, types of pesticide application, and crops/animals raised. In coding the data, these responses were changed to yes/no variables as a result of being marked or unmarked. When using these data, it is important to consider that the appropriate reference group may be those that did not do anything rather than those that did not do one particular activity.

When comparing the impact of pesticide application practices, for example, there are at least eight options that subjects may report using. If you are interested in the impact of using a backpack sprayer, the appropriate comparison group is probably those that did not apply any pesticides rather than those that did not use a backpack sprayer. These subjects may have used an air blast or mist fogger or some other risky form of application and keeping them in the comparison group may attenuate any observed effect.

Other ways of selecting reference groups may be appropriate as well.

Questions to which this consideration applies are Enrollment 9, 16, 17, 31. There are others in the take-home questionnaires as well.

5.4 Pesticide Grouping Analyses

There are four **functional classes** of pesticides included in the questionnaire: herbicides, insecticides, fungicides and fumigants. Table 5-4 lists the relevant question numbers for each class of pesticide in the questionnaires.

Table 5-4. Questions asking about usage of functional classes of pesticides in each questionnaire

Functional Class	Variable	Questionnaire			
		Enrollment	Farmer	Commercial	Spouse
Herbicide	A_EV_HRB	11 a-j; 12 n-u	19 a-h	11 a-h	9 b-s
Insecticide	A_EV_INS	11 k-s; 12 a-m	20 a-m	12 a-m	10A b-n; 10B a-f; 11 b-d
Fungicide	A_EV_FNG	11 u-v; 12 v-y	21 a-d	13 a-d	13 b-g
Fumigant	A_EV_FUM	11 t; 12 z,aa,bb	22 a-c	14a-c	12A b-c; 12B a-b

We have constructed a set of four indicator variables (flags) for applicators and a parallel set of four indicator variables for spouses. These have the value 1 if the respondent indicated that he or she had used one of the chemicals in the functional class and the value 0 if the respondent did not indicate use of any of the chemicals in the class

In addition to functional classes, there can be further subdivision of pesticides by chemical structure or mechanism of action. To date, we have created variables for seven **chemical classes** of pesticides: organochlorine insecticides, organophosphate insecticides, organothiophosphate insecticides (actually a subset of organophosphate insecticides,), carbamate pesticides, phenoxy herbicides, and triazine herbicides. These classes incorporate the pesticides included in the Enrollment and Spouse questionnaires. They also include the pesticides listed in Farmer Questionnaire questions 19-22 and in the Commercial Questionnaire questions 11-14. The chemical classes are defined as:

- Organochlorine insecticides (variable A_EV_OCH):
  - Aldrin;
  - Chlordane;
  - DDT;
  - Dieldrin;
  - Heptachlor;
  - Lindane; and
  - Toxaphene.

- Organophosphate insecticides (variable A_EV_OPH):
  - Chlorpyrifos;
  - Coumaphos;
  - Diazinon;
  - Dichlorvos/DDVP;
  - Fonofos;
  - Malathion;
  - Parathion;
  - Phorate;
  - Terbufos; and
  - Trichlorfon.
- Organothiophosphate insecticides (variable A_EV_OTH):
  - Chlorpyrifos;
  - Coumaphos;
  - Diazinon;
  - Fonofos;
  - Phorate; and
  - Terbufos.
- Carbamate pesticides (includes insecticides and one fungicide -- variable A_EV_CAR):
  - Aldicarb;
  - Carbofuran;
  - Carbaryl; and
  - Benomyl.
- Phenoxy herbicides (variable A_EV_PNX):
  - 2,4-D;
  - 2,4,5-T P; and

- 2,4,5-T.

- Triazine herbicides (variable A_EV_THI):
  - Atrazine;
  - Cyanazine; and
  - Metribuzin.

Note that the pesticides contained in the checklists in Question 24 of the Farmer Questionnaire, and Question 16 of the Commercial Applicator Questionnaire are not included in either the derived functional classes or in the derived chemical classes. Since these questions were not asked of all respondents and may contain less complete information, these variables were not included in the summary variables. It may, however, be appropriate for some analytic purposes to consider expanding the predefined chemical classes to include these variables when determining how to classify a specific respondent.

For a variety of reasons, it may be warranted to create groups of pesticides rather than to conduct pesticide-specific analyses. When combining pesticides into groups it is important to take the following points into consideration:

- The duration-of-use categories differ between the Enrollment and the take-home questionnaires. In the Enrollment Questionnaire, the highest category is greater than 30 years; in the take-home questionnaire it is greater than 20 years.
- The frequency-of-application categories are the same for herbicides and insecticides across the Enrollment and the take-home questionnaires but differ for fumigants and fungicides. This makes it difficult to combine over categories.
- When creating summary variables for frequency of application, it is possible to create values for annual pesticide use that exceed 365 days. A variety of approaches are available to standardize these variables.

A functional class flag or chemical class flag was set to "Yes" for a spouse if he or she responded "Yes" to the usage question for any of the pesticides in the list for that class (see Spouse Questionnaire questions 9-13). Similarly, a functional class flag or chemical class flag was set to "Yes" for an applicator if he or she responded "Yes" to the usage question for any of the pesticides in the list for that class in either the Enrollment Questionnaire or in a take-home questionnaire – either the Farmer Questionnaire or the Commercial Applicator Questionnaire.[2]

[2] No applicator has data in the file from both the Farmer Questionnaire and the Commercial Applicator Questionnaire.

Note that the 28 pesticides listed in Question 12 of the Enrollment Questionnaire were repeated in both the Farmer Application and Commercial Applicator Questionnaires. The reason for this is that, while only the basic Yes/No usage question was asked in the Enrollment Questionnaire, the full set of usage (Yes/No), duration, frequency, and decade of first use questions were asked in the take-home questionnaires. For purposes of creating the chemical class flags, a response of "Yes" in either the Enrollment Questionnaire or one of the take-home questionnaires was taken as "Yes" even if the applicator responded "No" in another questionnaire.

Since take-home questionnaires were completed by enrolled applicators, those who completed both a take-home questionnaire and the Enrollment Questionnaire (approximately 40% of the applicators) had more opportunity to have "Yes" responses recorded. As shown in Table 5-5 and Table 5-6, the increase in the number of applicators flagged as exposed is, on the average, small.

Table 5-5. Contribution of Enrollment and Take-Home Questionnaires to the Number of Applicators Classified as Exposed to Each Class of Pesticides—Counts

		Number of Cohort Exposed to Class		
	Number with at Least One Codable Data Point	Based on Enrollment Questionnaire Alone	Increment from Take-Home Questionnaire	Based on both Enrollment and Take-Home Questionnaires
Functional classes:				
Herbicide	56,599	53,299	543	53,842
Insecticide	56,567	49,928	1,158	51,086
Fungicide	56,441	19,080	1,109	20,189
Fumigant	56,390	12,356	696	13,052
Chemical classes:				
Organochlorine insecticides	53,980	25,117	1,762	26,879
Organophosphate insecticides	56,540	47,196	1,127	48,323
Organothiophosphate insecticides	56,511	41,216	872	42,088
Carbamate pesticides	55,026	34,435	1,331	35,766
Phenoxy herbicides	56,485	42,104	245	42,349
Triazine herbicides	56,544	41,677	220	41,897

Table 5-6. Contribution of Enrollment and Take-Home Questionnaires to the Number of Applicators Classified as Exposed to Each Class of Pesticides—Percents

		Percent of Cohort Exposed to Class		
	Number with at Least One Codable Data Point	Based on Enrollment Questionnaire Alone	Increment from Take-Home Questionnaire	Based on both Enrollment and Take-Home Questionnaires
Functional classes:				
Herbicide	56,599	94	1	95
Insecticide	56,567	88	2	90
Fungicide	56,441	34	2	36
Fumigant	56,390	22	1	23
Chemical classes:				
Organochlorine insecticides	53,980	47	3	50
Organophosphate insecticides	56,540	83	2	85
Organothiophosphate insecticides	56,511	73	2	74
Carbamate pesticides	55,026	63	2	65
Phenoxy herbicides	56,485	75	0	75
Triazine herbicides	56,544	74	0	74

As with other groups of pesticides, when considering duration and frequency of exposure to chemical classes, it is important to take the following points into consideration:

- The duration-of-use categories differ between the enrollment and the take-home questionnaires. In the enrollment questionnaire, the highest category is greater than 30 years; in the take-home questionnaire it is greater than 20 years.

- The frequency-of-application categories are the same for herbicides and insecticides across the enrollment and the take-home questionnaires but differ for fumigants and fungicides. This makes it difficult to combine over categories.

- When creating summary variables for frequency of application, it is possible to create values for annual pesticide use that exceed 365 days. A variety of approaches are available to standardize these variables.

As the functional and chemical class flag groups are defined only in terms of the ever/never use responses, they are not affected by factors related to the duration, frequency, or intensity of use.

5.5 Use of Midpoint Codes

In each of the questionnaires there are questions that ask the respondent to select a response representing a range. For example, the choices for years of pesticide application in the Enrollment Questionnaire are “1 year or less,” “2-5 years,” 6-10 years,” “11-20 years,” “21-30 years,” and “More than 30 years.” Rather than code the responses in the database for these types of questions with the somewhat arbitrary set of sequential numbers, “1,” “2,” “3”, “4,” “5,” and “6” as is customarily done, we have chosen to code them with values representing the “midpoints” of each range. In this case, the codes used are “1,” “3.5,” “8,” “15.5,” “25.5,” and “35.” While there can be little argument about most of these values, the choice of the lowest and highest in each case represents a judgment by a committee of analysts about what values will be most useful for typical analyses. It is quite likely that different values will be appropriate for some analyses.

Analysts are free to select any values for the midpoints that are, in their judgment, appropriate for their analyses. Since the values in the file have proved to be useful in a number of analyses to date, we recommend their use unless there is a specific reason to change them. The values in the file can be used as shipped in procedures such as SAS’s PROC LOGIT.

Note that, in some cases, when a zero response is implied by a preliminary question, the value for a range was set to “0” even though that is not one of the choices. Thus, when a farmer indicated that he had not used atrazine, the reply to “How many years did you personally mix or apply this pesticide?” was set to “0” rather than missing.

5.6 Analysis of Fungicide Duration and Frequency Responses

In the Enrollment and applicator questionnaires, with a few exceptions, participants were asked about their fungicide usage in a similar way to other types of pesticides. For duration of use of fungicides, subjects were given the option of reporting “already applied to seed” without providing a number of years. Similarly for frequency of use, participants were given the options of “pre-applied to seed” and “none” for the number of days per year applied.

Since all these activities will result in some fungicide exposure, or at least in greater fungicide exposure than among those who did not use fungicides at all, in the analytic dataset we have assigned a value of 0.1 for pre-applied to seed for duration and frequency and have assigned

a value of ".N" for those who stated no days of application. These are temporary values until we can gather experts to create consensus estimates regarding fungicide exposure. It is anticipated that these consensus estimates will be available by the next release of the data. To complicate matters further, farmers who used only treated seed may not have reported using fungicides. This will also be explored for the next release of the data.

5.7 Pesticides Used in Combination with Other Pesticides

The questionnaires for both applicators and spouses ask for information regarding 50 separate pesticides or pesticides used in combination. As a result of pesticides asked in combination, there are 52 unique pesticides included on the questionnaire. The pesticides asked in combination are:

- Maneb/Mancozeb;
- Permethrin or other pyrethroids; and
- Carbon Tetrachloride/Carbon Disulfide (80/20 mix).

For Permethrin, an insecticide used both on crops and animals, pesticide usage is asked separately for crop and animal usage. No other pesticides are asked separately for crop and animal application.

5.8 Unusual Values

It is common in large-scale surveys for some percentage of the respondents to provide some inconsistent or improbable answers. This is particularly true in questionnaires such as those used in this study which were completed by respondents with no opportunity to ask study staff about the meanings of terminology or phrasing.

We have performed extensive reviews of the data in order to locate and resolve seemingly inconsistent data. We also reviewed reported values which were unusually large. In some cases, we were able to review the optically scanned forms and determine that the response recorded during scanning was not what the respondent intended. This occurred occasionally because of incomplete erasures or other problems with the source document. In addition, the field workers in Iowa and North Carolina who, during a followup phase of the study, have been

telephoning to respondent households, have independently verified all the birth dates of those contacted. Birth dates are available for all but two private applicators and four spouses. In addition one private applicator provided a day and year of birth, but not month. The birth month was set to six for this person. Four other private applicators provided birth month and years, but not days. Their days of birth were each set to 15. In each of these cases the imputation flags A_BIRTH_I (in the applicator files) and S_BIRTH_I (in the spouse file) were set to 1 or 2 to record the type of change.

The most extensive editing efforts involving physically reviewing the source documents were performed on the private applicator data and very limited checks have been performed on data from commercial applicators. A moderate amount of checking was performed on the data relating to spouses.

Questions relating to medical conditions were asked in the Enrollment, Farmer, Commercial Applicator, and Spouse Questionnaires. Editing performed on them and considerations in their analysis are discussed in the next section.

Any inconsistencies remaining after these reviews reflect the actual responses of the pesticide applicators and their spouses. While some of the remaining responses may be inconsistent within a questionnaire or across questionnaires, they do reflect what the respondents entered. A list of responses remaining in the data that fail the consistency and plausibility checks is available to researchers on request. Researchers encountering apparently anomalous data values are encouraged to report them to the Agricultural Health Study by sending an e-mail message to the Agricultural Health Study Coordinating Center at CoordinatingCenter@aghealth.org.

5.9 Medical Condition Variables

Questions relating to medical conditions were asked in the Enrollment, Farmer Applicator, Commercial Applicator, and Spouse Questionnaires. The question number and number of conditions asked about are displayed in Table 5-7.

Table 5-7. Location and Number of Medical Condition Questions

Questionnaire	Question Number	Number of Medical Conditions Listed
Enrollment	28	16
Farmer (Private Applicator)	87	41
Commercial Applicator	75	41
Spouse	105	49

The three take-home questionnaires (Farmer, Commercial Applicator, and Spouse) also asked for the age range at which a condition was diagnosed. The ranges were Younger than 20, 20-39, 40-59, and 60 or older. Some respondents answered these questions in a clearly erroneous fashion. For example, some respondents answered "Yes" to all of the medical conditions on a questionnaire. Another problem that appeared a number of times occurred when a participant responded "No" to the first several medical conditions, then responded "Yes" to a condition and moved to the right of the page to enter the age of diagnosis. Subsequent responses were either a series of "Yes" responses (with no age of diagnosis) or a blank response to the Yes/No question but a positive response to the leftmost age of diagnosis (Younger than 20).

An initial automated edit set the Yes/No responses to "Yes" when an age of diagnosis was supplied. When it became obvious that this was producing an unrealistically high number of "Yes" responses, we removed this edit and carefully examined all Farmer and Spouse Questionnaires with 10 or more "Yes" responses to the medical condition question or 5 or more "Yes" responses in a row. Where it was practical to determine the intent of the respondent, the entries in the data files were corrected for these questionnaires. No similar reviews have been performed for the Enrollment Questionnaire data or the Commercial Applicator Questionnaire data.

We are aware of additional anomalies in the responses to the medical condition variables in a small percentage of the records and may be able to correct some of them in future updates to these data files. For instance, while 99.5 percent of the respondents indicated that they had fewer than four of the listed medical conditions on the Enrollment Questionnaire, and 99.5 percent of the respondents to the Farmer, Commercial Applicator, and Spouse Questionnaires had fewer than eight medical conditions, there were 0.5 percent of the respondents to each questionnaire who gave greater numbers of positive responses. While some of these responses are undoubtedly accurate reflections of the respondents' medical conditions, some are certainly inaccurate. Analysts using these data may want to apply the following edit rule to adjust for such anomalous responses:

If a respondent replied “Yes” to 16 or more medical conditions, change those responses to “No” and change any “No” responses to “Yes.”

After carefully reviewing the response patterns to the 16 medical condition questions in the Enrollment Questionnaire, we determined to interchange “Yes” and “No” responses to the medical condition questions for people who responded “Yes” to 10 or more of the 16 conditions. This edit affected 24 records.

As part of the preparations for the 2003 release of the Phase I data, we investigated alternate cutoff points for correcting responses to the medical condition questions in the Enrollment Questionnaire. Using a lower number of “Yes” responses did not improve the plausibility of any of the response patterns that would be affected. We also decided to set all of the medical condition responses for two IDs (616011 and 619465) to missing because they clearly alternated their “Yes” and “No” responses to the list of medical conditions.

We also applied an analytic edit to medical condition response patterns consisting of one or more “Yes” responses, no “No” responses, and some missing responses. For this pattern, we changed the missing responses to “No.” An archival version of the data is available in case a researcher needs to review data prior as it stood prior to this change. Note that response patterns containing at least one “Yes” and one “No” response did not have missing responses changed. They remain missing.

5.10 Interpretation of Missing Data Patterns

Analysts should be aware of four types of missing data:

- Missing values;
- Patterns of missing data associated with questionnaires completed by phone;
- Questionnaires collected during a reliability study; and
- Respondents skipping a page or stopping.

Each of these types of missing data is discussed separately below.

5.10.1 Missing Value Codes

The majority of the fields in the Phase I analysis files were converted from character to numeric. During the conversion process, any nonnumeric values automatically convert to missing, which is represented by a period. SAS has the ability to represent different categories of missing data by using a period followed by a letter. In addition to the SAS default missing value code, three other missing values codes appear in the Phase I datasets to represent special conditions that may be useful to researchers to distinguish:

1. Many of the fields in the optically scanned questionnaires had multiple responses. These were originally coded with an asterisk (*) by the scanning program. Since SAS does not accept asterisks in numeric fields, these multiple responses codes were represented in the analysis files with the missing value code ".M".

2. Approximately 9 percent of the responses to the Spouse Questionnaire items regarding the highest level of schooling completed contained a letter instead of the valid questionnaire responses of 1-9. These invalid codes were set to ".U". If these values are needed for future analyses, the hard-copy questionnaires should be reviewed to determine if a valid response codes can be assigned.

3. The Pesticide Use section of the Enrollment, Farmer Applicator, and Commercial Applicator Questionnaires contains a series of questions about fungicides used. When a respondent indicates that he or she used a fungicide, the column next to it which is headed by the question "In an average year when you personally used this pesticide, how many days did you use it?" contains the response choices "Pre-applied to seed," "None," "1 day," "2-5 days," "3-9 days," and so on. The response "None" is unexpected, since the only time a respondent should be marking this section of the questionnaire is after having indicated that he or she personally mixed or applied the fungicide. Since the "None" response is more like a missing value than a true zero, it has been set to ".N" in the data files. Users who want to treat this response as zero should recode the ".N" values to zero. The only variables with values of ".N" are those relating to the number of days respondents used fungicides.

5.10.2 Questionnaires Completed by Phone

As discussed in Section 2.2, a substantial number of questionnaires were completed by telephone. In case people had only a limited amount of time to respond to the questionnaires, the questions were grouped into modules and asked in priority order: 1. Application, 2. Health, 3. Lifestyle. Analysts should be aware that some patterns of missing data may be related to time constraints and the order in which the questions were asked.

5.10.3 Questionnaires Collected during Reliability Study

A Followup Questionnaire, which contains a subset of the questions in the Enrollment Questionnaire, was also administered to 2,895 people in Iowa as part of a reliability study (Blair, et al., 2002) . It was administered in the same way as the original Enrollment Questionnaire; that is, it was completed at county agricultural extension office when the applicator arrived for a pesticide training course. Although it was intended that the respondents to this questionnaire were people who had completed an Enrollment Questionnaire one year previously, 344 new respondents completed the Followup Questionnaire but not the Enrollment Questionnaire. The Followup Questionnaire data from these new respondents, are included with the Enrollment Questionnaire data. The data from the repeat respondents are not included in these files.

Because the respondents who were enrolled using the Followup Questionnaire rather than the full Enrollment Questionnaire did not have the opportunity to respond to many of the questions on the Enrollment Questionnaire, their responses to these questions have been set to missing. Some of these respondents were subsequently called as part of the process to collect key missing data (see Section 2.4). While the responses from this subset of the Agricultural Health Study cohort are valid and useful, analysts need to recognize that most of their missing data is an artifact of the data collection methodology and does not represent a desire by the respondents to avoid answering some of the questions.

5.10.4 Respondent Missed a Page or Stopped

In some instances, it is clear that a respondent simply skipped one or more pages of a questionnaire. Researchers analyzing patterns of missing data may need to take this fact into account and check whether all the items physically grouped onto the same page have missing values or values indicating that the variable was not marked.

Similarly, some respondents completed part of a questionnaire and simply stopped filling it out. Researchers analyzing missing data patterns may also want to check for this type of pattern.

5.11 SIC/SOC Coding

Responses to questions in the Farmer Questionnaire and Spouse Questionnaire regarding work off the farm were assigned Standard Industrial Classification (SIC) Codes and Standard Occupational Classification (SOC) Codes. Similar coding was **not** performed for the responses to these questions in the Commercial Applicator Questionnaire.

Although the SOC codes are reported as six-character codes, the SOC codes in the Private Applicator have been coded to the level of detail expressed in the first four characters. The fifth and sixth characters were added to allow researchers space to add additional specificity in the future without changing the format of the variables or adding a new variable. The fifth and sixth characters on the file are currently set to "00."

The values of the SOC codes in the Spouse File are also expressed in terms of 6 characters in order to maintain the same format as in the Applicator File. The fifth and sixth digits of the SOC codes in the Spouse File are also set to "00."

The textual responses to the nonfarm job questions are not stored on the file, but can be found in the appropriate verbatim response file (see Section 4.9).

5.12 Limitations and Uses of the Commercial Applicator File

The data in the Commercial Applicator File consist of the data for commercial applicators collected on the Enrollment Questionnaire and the data collected on the Commercial Applicator Questionnaire. Although the file structure is identical to that of the Private Applicator File, analysts should be aware that there are some subtle differences between the private applicator data and the commercial applicator data.

Because the private and commercial applicators are different people, for some analyses it may be appropriate to concatenate the data in the two corresponding data files. This has the effect of increasing the number of records in the analysis, but this gain may be illusory.

In most cases, the identical questions are presented in the Farmer Applicator and Commercial Applicator Questionnaires. There are a few questions relating to the distance of fields from the applicator's home or well that were not asked of commercial applicators. Since the Private and Commercial Applicator Files have the same structure, these variables all have missing values on the Commercial Applicator File. In addition, some questions were asked slightly

differently on the two take-home questionnaires. Investigators will need to determine whether these differences have analytic implications for the analyses they are conducting. The questions with wording differences are shown in Table 5-8.

Table 5-8. Questions in the Farmer Applicator Questionnaire and Commercial Applicator Questionnaire with Wording Differences

Farmer Questionnaire Item	Commercial Questionnaire Item
Q11. What application methods do you generally use when you apply crop insecticides? (Mark all that apply.)	Q3. What application methods do you generally use when you apply crop, nursery, lawn and garden insecticides? (Mark all that apply.)
Q16. When you personally mix crop insecticides, what additives do you generally use? (Mark all that apply.)	Q8. When you personally mix crop, nursery, lawn and garden insecticides, what additives do you generally use? (Mark all that apply.)
Q51. Did you ever have a job off a farm?	Q39. Did you ever have a job other than as a commercial pesticide applicator?
Q52. For the nonfarm job you held the longest, what was your job?	Q40. For the job you held the longest (other than as a commercial pesticide applicator), what was your job?
Q54. For the nonfarm job you held the longest, which of the following were you exposed to? (Mark all that apply)	Q42. For the job you held the longest (other than as a commercial pesticide applicator), which of the following were you exposed to? (Mark all that apply)

Commercial applicators were not given Spouse Questionnaires to take home, so analyses requiring responses to both the Commercial Applicator Questionnaire and Spouse Questionnaire cannot be conducted. Since some Commercial Applicators are female, it is possible to merge records from the Female and Family Health File with those from the Commercial Applicator File. When performing such merges, analysts need to be careful to check the gender variable in the Commercial Applicator File to ensure that it indicates that the applicator is female.

5.13 Order and Naming of Medical Condition Variables

A series of questions is asked in the Enrollment, Farmer Applicator, Commercial Applicator, and Spouse Questionnaires relating to medical conditions with which respondents have been diagnosed. The three take-home questionnaires also collect the age of diagnosis. The list of medical conditions varies across the questionnaires. Even when the same conditions are included in a pair of questionnaires, the order of the questions sometimes varies. In order to make programming analyses easier, a consistent naming convention was established for the medical conditions. This was done by making an alphabetic list of all the medical conditions listed in the

four questionnaires and then naming the variables A_MEDCOND1, A_MEDCOND2, etc. in the Private Applicator File and the Commercial Applicator File. The corresponding variables in the Spouse File and the Supplemental Spouse File are named S_MEDCOND1, S_MEDCOND2, etc. The diagnosis age variables in the applicator files have the form A_AGECOND1, A_AGECOND2, etc., and are numbered in the same order as the corresponding medical conditions.

When combining the medical conditions from Enrollment Questionnaire Question 28, Farmer Questionnaire Question 87, Commercial Questionnaire Question 75 and Spouse Questionnaire Question 105, one finds 57 unique categories. Even though only a subset of the 57 categories exists on each file, for programming ease in setting up arrays, we represented them all on the final analysis files and set the value to missing if the question was not asked on a specific questionnaire. For example, S_MEDCOND1 represents responses to a question on Alzheimer's disease from the Spouse Questionnaire. Since there is no corresponding question in the Enrollment, Farmer or Commercial Questionnaires, a placing holding variable A_MEDCOND1 was placed in the applicator files and always has a value of missing (“.”). Similarly, A_MEDCOND5 represents a variable in the Enrollment Questionnaire asking about asthma. Since the age of diagnosis is not asked in the Enrollment Questionnaire, the variable A_AGECOND5 was created as a placeholder in arrays, but always has the value of missing (“.”).

For the convenience of people who have been using preliminary versions of the data files, a crosswalk between the old and new variable names is provided (see the next section).

Two identical medical condition questions exist in the enrollment and farmer/commercial data. They are the questions asking about diabetes and Parkinson's disease. Since Enrollment and Farmer Questionnaire data were combined into one data file data (as were Enrollment and Commercial Questionnaire data), both sets of responses to these questions are retained. For diabetes, A_MEDCOND16 contains the medical conditions from the farmer/commercial data and A_MEDCOND16E contains the medical conditions from enrollment. For Parkinson's disease, A_MEDCOND44 contains the medical conditions from the farmer/commercial data and A_MEDCOND44E contains the medical conditions from enrollment.

While there is a great deal of similarity between the two sets of responses, they are not identical. For instance, 187 farmers (0.36%) changed from “Yes” to “No” or vice versa when responding to the diabetes question, and 781 (1.6%) left the response blank in one of the questionnaires and responded “Yes” in the other. Similarly, 12 farmers (0.02%) changed their

responses from “Yes” to “No” or vice versa on the Parkinson’s disease question and 27 (0.05%) left the response blank in one of the questionnaires, but answered “Yes” in the other. Similar response patterns (ranging from 0 to 0.55%) for these two pairs of questions exist for commercial applicators. Thus, researchers analyzing these data need to decide how to treat these differential responses.

There are four sets of medical condition questions that are similar across questionnaires, but have been treated as different variables because of differences in the question wording. These are the variables asking about asthma, depression, kidney disease, and pneumonia. Researchers may want to treat the responses to some of these questions as comparable. They should review the question wording in the different questionnaires to determine what is appropriate for each specific analysis.

5.14 Applicator-Spouse Variable Crosswalk

As an aid to programmers who have used preliminary versions of these data files, we have included as part of the documentation an Excel spreadsheet called Applicator_Spouse_Crosswalk.xls. This spreadsheet lists all of the variables from the Farmer, Commercial and Spouse Questionnaires that correspond to variables in the final Private Applicator Analysis File.

Names in the final file have the same naming convention as the original Phase I data files, but the first letter is now an “A” instead of “E,” “F,” or “C”. For example ESCHOOL is now ASCHOOL. The most extensive naming convention changes occurred in the pesticides and medical conditions. The pesticides were reordered to match the order in the Spouse File. The naming conventions of the variables were expanded to a longer name. For example: A_HERBICIDE_CD1 was originally EHRBCD1 in the preliminary Enrollment Questionnaire File. The same variable in the Spouse File is called S_HERBICIDE_CD1 and was originally called SHRBCD1. The medical conditions were also reordered in alphabetic order by condition name. The crosswalk for these variables is also included in the spreadsheet.

5.15 Use of Supplemental SAS Format and Attribute Statements

Use of a common set of format values can increase consistency among AHS reports and reduce programming time. To facilitate analyses, we have included a set of tested SAS

statements in the directory FormatsAndAttributeSASCode, which programmers can easily add to their own programs.

The file "ProcFormat_SAS.txt" contains a set of SAS PROC FORMAT values that are commonly used in our AHS analyses. These SAS statements can be inserted into a PROC FORMAT and are ordered in the following manner:

- Generic formats (formats appearing in multiple Questionnaires);
- Enrollment formats (formats appearing in the Enrollment Questionnaire);
- Farmer formats (formats appearing in the Farmer Questionnaire);
- Spouse formats (formats appearing in the Spouse Questionnaire); and
- Female and Family Health formats (formats appearing in the Female and Family Health Questionnaire).

In order to use these formats, it is necessary to refer to them in SAS data steps or procedures. We have created three files which can be inserted into SAS data steps to associate each questionnaire variable with the appropriate format:

- Appl_Format_Statement_SAS.txt – The statements in this file can be used in SAS data steps referring to either the Private Applicator File or the Commercial Applicator File.
- Spouse_Format_Statement_SAS.txt – The statements in this file can be used in SAS data steps referring to either the Spouse File or the Supplemental Spouse File.
- FFH_Format_Statement_SAS.txt – The statements in this file can be used in SAS data steps referring to the Female and Family Health File.

All three of these files refer to formats defined in ProcFormat_SAS.txt.

In order to facilitate analyses of the Mortality Data File, we have included format statements for both ICD-9 and ICD-10 cause of death codes. These are in files ICD9_FORMAT.TXT and ICD10_FORMAT.TXT, respectively. The formats translate ICD codes into medical descriptions. For example, the first four ICD-10 values are translated as follows:

```
"A00"  = "Cholera"
"A000" = "Cholera due to Vibrio cholerae 01, biovar cholerae"
"A001" = "Cholera due to Vibrio cholerae 01, biovar el tor"
"A009" = "Cholera, unspecified"
```

Use the variable ICD_VERSION in the Mortality Data File to determine whether to use the ICD-9 or the ICD-10 formats.

It is often useful to have SAS reports include the question number and a short description of the question instead of, or in addition to, the variable name. We have created three files of SAS attribute statements which assign appropriate SAS labels to each questionnaire variable. These statements also order the variables in the order they appear in the questionnaires, with slight modifications to allow easier use of SAS array processing capabilities:

- Attrib_PrivateApplicator_SAS.txt – The statements in this file can be used in SAS data steps referring to either the Private Applicator File or the Commercial Applicator File.
- Attrib_Spouse_SAS.txt – The statements in this file can be used in SAS data steps referring to either the Spouse File or the Supplemental Spouse File.
- Attrib_FFH_SAS.txt – The statements in this file can be used in SAS data steps referring to the Female and Family Health File.

6. REFERENCES

The following references contain a 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/facts.html#publications.

The publications listed below have been divided into four major sections: Methods, Exposure Assessment, Health Outcomes, and Diet.

6.1 Methods

Alavanja, M. C., Akland, G., Baird, D., Blair, A., Bond, A., Dosemeci, M., Kamel, F., Lewis, R., Lubin, J., Lynch, C., & et al. (1994). Cancer and noncancer risk to women in agriculture and pest control: the Agricultural Health Study. *J Occup Med*, 36(11), 1247-1250.

Alavanja, M. C., Sandler, D. P., McDonnell, C. J., Lynch, C. F., Pennybacker, M., Zahm, S. H., Mage, D. T., Steen, W. C., Wintersteen, W., & Blair, A. (1999). Characteristics of pesticide use in a pesticide applicator cohort: the Agricultural Health Study. *Environ Res*, 80(2 Pt 1), 172-179.

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. Accuracy of self-reported pesticide use duration information from licensed pesticide applicators in the Agricultural Health Study. *J Expo Anal Environ Epidemiol* (in press).

Sandler D. P., Wegienka G., Baird D. D., Gladen B. C., Rowland A. S., Alavanja M. C. R., Logsden-Sackett N., Pierce J., Lynch C. F., Knott C., Blair A. Assessing response bias: examples using data from young women in the Agricultural Health Study. (Submitted)

Tarone, R. E., Alavanja, M. C., Zahm, S. H., Lubin, J. H., Sandler, D. P., McMaster, S. B., Rothman, N., & Blair, A. (1997). The Agricultural Health Study: factors affecting completion and return of self-administered questionnaires in a large prospective cohort study of pesticide applicators. *Am J Ind Med*, 31(2), 233-242.

6.2 Exposure Assessment

Coble J., Hoppin J. A., Engel L., Elci O. C., Dosemeci M., Lynch C.F., Alavanja M. (2002). Prevalence of exposure to solvents, metals, grain dust and other hazards among farmers in the Agricultural Health Study.

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.*

Gladen, B. C., Sandler, D. P., Zahm, S. H., Kamel, F., Rowland, A. S., & Alavanja, M. C. (1998). Exposure opportunities of families of farmer pesticide applicators. *Am J Ind Med*, 34(6), 581-587.

6.2.1 High Pesticide Exposure Events

Alavanja, M. C., Sandler, D. P., McDonnell, C. J., Lynch, C. F., Pennybacker, M., Zahm, S. H., Lubin, J., Mage, D., Steen, W. C., Wintersteen, W., & Blair, A. (1998). Factors associated with self-reported, pesticide-related visits to health care providers in the Agricultural Health Study. *Environ Health Perspect*, 106(7), 415-420.

Alavanja, M. C., Sandler, D. P., McDonnell, C. J., Mage, D. T., Kross, B. C., Rowland, A. S., & Blair, A. (1999). Characteristics of persons who self-reported a high pesticide exposure event in the Agricultural Health Study. *Environ Res*, 80(2 Pt 1), 180-186.

Alavanja, M. C., Sprince N. L., Oliver E., Whitten P., Lynch C., Gillette P. P., Logsden-Sacket N., Zwerling C. (2001).A nested case-control analysis of high pesticide exposure events from the Agricultural Health Study. *Am J Ind Med*, 39(6):557-63.

Keim, S. A. & Alavanja, M. C. (2001). Pesticide use by persons who reported a high pesticide exposure event in the Agricultural Health Study. *Environ Res*, 85(3), 256-259.

Mage, D. T., Alavanja, M. C., Sandler, D. P., McDonnell, C. J., Kross, B., Rowland, A., & Blair, A. (2000). A model for predicting the frequency of high pesticide exposure events in the Agricultural Health Study. *Environ Res*, 83(1), 67-71.

6.2.2 Environmental Measures

Brock, J. W., Melnyk, L. J., Caudill, S. P., Needham, L. L., & Bond, A. E. (1998). Serum levels of several organochlorine pesticides in farmers correspond with dietary exposure and local use history. *Toxicol Ind Health*, 14(1-2), 275-289.

Geno, P. W., Camann, D. E., Harding, H. J., Villalobos, K., & Lewis, R. G. (1996). Handwipe sampling and analysis procedure for the measurement of dermal contact with pesticides. *Arch Environ Contam Toxicol*, 30(1), 132-138.

Melnyk, L. J., Berry, M. R., & Sheldon, L. S. (1997). Dietary exposure from pesticide application on farms in the Agricultural Health Pilot Study. *J Expo Anal Environ Epidemiol*, 7(1), 61-80.

Shealy, D. B., Barr, J. R., Ashley, D. L., Patterson, D. G., Jr., Camann, D. E., & Bond, A. E. (1997). Correlation of environmental carbaryl measurements with serum and urinary 1-naphthol measurements in a farmer applicator and his family. *Environ Health Perspect*, 105(5), 510-513.

Shealy, D., Bonin, M., Wooten, J., Ashley, D., Needham, L., & Bond, A. (1996). Application of an improved method for the analysis of pesticides and their metabolites in the urine of farmer applicators and their families. *Environmental International*, 22(6), 661-675.

6.3 Health Outcomes

Hoppin J. A., Umbach D. M., London S. J., Alavanja M. C. R., Sandler D. P. (2002) Chemical predictors of wheeze among farmer pesticide applicators in the Agricultural Health Study. Am J Resp Crit Care Med, 165(5), 683-9.

Hoppin J. A., Umbach D. M., London S. J., Alavanja M. C. R., Sandler D. P. (2003). Animal production and wheeze in the Agricultural Health Study: Interactions with asthma, atopy, and smoking.

Kamel, F., Boyes, W. K., Gladen, B. C., Rowland, A. S., Alavanja, M. C., Blair, A., & Sandler, D. P. (2000). Retinal degeneration in licensed pesticide applicators. *Am J Ind Med*, 37(6), 618-628.

Martin S. A., Sandler D. P., Harlow S. D., Shore D. L., Rowland A. S., Alavanja M. C. R. Pesticide exposure and pesticide-related symptoms among black farmers in the Agricultural Health Study. (2002) *American Journal of Industrial Medicine*, 41:202-209.

Martin S. A., Harlow S. D., Sowers M.F., Longnecker M. P., Garabrant D., Shore D. L., Sandler D. P. (2002). DDT metabolites and androgens in African-American farmers. Epidemiology, 13(4): 454-8.

Rowland A. S., Baird D. D., Long S., Wegienka G., Harlow S. D., Alavanja A, Sandler D. P. (2002) Influence of Medical conditions and lifestyle factors on the menstrual cycle. *Epidemiology*.

Sprince N. L., Zwerling C., Lynch C. F., Whitten, P.S., Thu K., Logsden-Sackett N., Burmeister L. F., Sandler D. P., Alavanja M. C. R. (2003). Risk factors for agricultural injury: a case-control analysis of Iowa farmers in the Agricultural Health Study. Journal of Agricultural Safety and Health.

6.4 Diet

Keating, G. A., Sinha, R., Layton, D., Salmon, C. P., Knize, M. G., Bogen, K. T., Lynch, C. F., & Alavanja, M. (2000). Comparison of heterocyclic amine levels in home-cooked meats with exposure indicators (United States). *Cancer Causes Control*, 11(8), 731-739.

7. DATA FILE REQUESTS

Extracts of the data files will be provided to qualified researchers approved by the Agricultural Health Study Executive Committee. Refer to Agricultural Health Study Policy 2-4 for the procedures for submitting data file requests. A copy of Policy 2-4 and a copy of the AHS Request Form can be found on the project Web site at http://aghealth.nci.nih.gov/usersmanual/datarequest.html.

The AHS Request Form should be accompanied by a list of variables to be included on the requested data file. The variable checklist (P1_checklist.xls), which is available on the AHS Web site, should be used for this purpose. The checklist is an Excel spreadsheet, with a separate tab of the spreadsheet for each distinct data file. The variable names are in column B and the variable labels are in column C. Place an X in column A next to each variable being requested. Note that a number of linking and demographic variables are marked by default.

Three spreadsheets are available by request to assist users in choosing variables for data requests. They are called Applicator_vars.xls, Spouse_vars.xls, and FFH_vars.xls. The first four columns contain information that will help users sort variables in a specific order. The last two columns contain the variable names and labels. Each spreadsheet has the same column structure:

- **Analysis File Order –** This column should be used in combination with column 2 or column 3. It orders the variables in Analysis file order with the questionnaire variables first followed by all of the derived variables. Since the applicator file blends Enrollment and Farmer data, the variables that cross the two questionnaires are mixed together.

- **New Variable Order within Question** – Some of the questions have multiple variables for responses. For the most part the variable order in the files is the same as their order in the questionnaire. An example of a set of questions where variable order changed is the medical condition questions in Farmer questionnaire question 87 and Spouse questionnaire question 105, which were reordered, in alphabetic order by condition. When used in combination with **Analysis File Order** which is the high ordering sort variable, the medical conditions will be in the new alphabetic order.

- **Old Variable Order within Question** – When used in combination with **Analysis File Order** which is the high order sort variable, the medical conditions will be in the original questionnaire order.

- **Variable Origin** – An additional sort field which contains "E" for enrollment, "F" for farmer, "S" for spouse and "D" for derived variables. It can be used to subset the variables to those associated with a specific questionnaire.

- **Variable:** The variable name in the analysis file.

- **Label:** The variable label in the analysis file.

Table 7-1. Sample from Applicator_vars.xls

Analysis File Order	New Variable Order within Question	Old Variable Order within Question	Variable Origin	Variable	Label
0	1	1	E	PARTID	Participant Identifier
1	1	1	E	A_ENROLLDT	E2: Date in which enrollment completed
1	2	2	E	ATMONTH	E2: Month in which enrollment completed
1	3	3	E	ATDAY	E2: Day in which enrollment completed
1	4	4	E	ATYEAR	E1: Year in which enrollment completed
200	1	1	D	A_AGE	Age

8. QUESTIONNAIRES

Five questionnaires were administered during Phase I of the study. These are:

- Enrollment Questionnaire – The questionnaire administered at enrollment.
- Farmer Applicator Questionnaire – The take-home questionnaire completed by private pesticide applicators.
- Commercial Applicator Questionnaire – The take-home questionnaire completed by commercial pesticide applicators.
- Spouse Questionnaire – The take-home questionnaire completed by spouses of private pesticide applicators.
- Female and Family Health Questionnaire – The take-home questionnaire completed by the female pesticide applicators and spouses of male applicators.

An annotated Portable Document Format (PDF) version of each of the questionnaires is available on the AHS web site at http://aghealth.nci.nih.gov/questionnaires.html. They are named "EnrollmentQuestionnaire.pdf," "FarmerQuestionnaire.pdf," "CommercialApplicatorQuestionnaire.pdf," "SpouseQuestionnaire.pdf." and "FemaleAndFamilyHealthQuestionnaire.pdf." The questionnaires look exactly like what the participants saw except that variable names have been added in square brackets and colored red in the annotated versions.

The Followup Questionnaire, which contains a subset of the questions in the Enrollment Questionnaire, is also included as part of the documentation. The annotated questionnaire is name "FollowUpQuestionnaire.pdf."

In addition, PDF files containing the following validation study questionnaires are included with the documentation:

- Neurologic and Immunologic Disease Questionnaire;
- Young Women's Health Questionnaire; and
- Women's Health Questionnaire.

9. CODEBOOKS

A detailed codebook for each of the data files has been included with the documentation files. As these codebooks are rather lengthy, they have been provided as separate files so that they can be selectively printed. The files are in rich text format and can be viewed by any word processor that recognizes rich text, in particular Microsoft Word for Windows and WordPerfect for Windows.

Each codebook lists all of the variables in its file and, for continuous variables provides, the range of valid values for the variable. For categorical variables, each valid code is listed along with its meaning. A detailed frequency distribution is provided for each categorical variable. Summarized frequency distributions are provided for continuous variables and ID variables.

The names of the codebook files are given in Table 9-1.

Table 9-1. Codebook file names

Codebook File Name	Description of Corresponding Data File
Main Questionnaire Files	
PrivateApplicator.rtf	Scanned responses from Farmer Applicator Questionnaire
CommercialApplicator.rtf	Scanned responses from Commercial Applicator Questionnaire
Spouse.rtf	Scanned responses from Spouse Questionnaire
FemaleAndFamilyHealth.rtf	Scanned responses from Female and Family Health Questionnaire
SupplementalSpouse.rtf	Scanned responses from Spouse Questionnaires that were completed by spouses who also enrolled separately as applicators
Verbatim Response Files Associated with Main Questionnaire Files	
PrivateApplicator_Verbatim.rtf	Verbatim responses from Farmer Applicator Questionnaire
CommercialApplicator_Verbatim.rtf	Verbatim responses from Commercial Applicator Questionnaire
Spouse_Verbatim.rtf	Verbatim responses from Spouse Questionnaire
FemaleAndFamilyHealth_Verbatim.rtf	Verbatim responses from Female and Family Health Questionnaire
SupplementalSpouse_Verbatim.rtf	Verbatim responses from Spouse Questionnaires that were completed by spouses who also enrolled separately as applicators
Validation Study Files	
NeurologicAndImmunologicDisesase.rtf	Neurologic and Immunologic Disease Study data
YoungWomensHealth_FFH.rtf	Young Women's Health Study data corresponding to questions in the Female and Family Health Questionnaire
YoungWomensHealth_SPO.rtf	Young Women's Health Study data corresponding to questions in the Spouse Questionnaire
WomensHealthStudy_FFH.rtf	Women's Health Study data corresponding to questions in the Female and Family Health Questionnaire
WomensHealthStudy_SPO.rtf	Women's Health Study data corresponding to questions in the Spouse Questionnaire

10. PHASE I ALGORITHM INTENSITY SCORES -- VERSION 2

This chapter presents the specification used to create the version 2 of the intensity scores and the intermediate variables used in the calculations.

10.1 Intensity Scores

Intensity scores is calculated as follows.

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

If the application did not personally apply, then the intensity score = 0.

NOTE: An applicator satisfies the "Did not personally apply" condition if the response to Question 10 "During your lifetime, have you ever personally mixed or applied any pesticides" is "No" (i.e., variable AMIXPERS = 0). An applicator will also meet the condition if AMIXPERS = 1, but APCTMIX = 0 (i.e., did not mix) and APCTAPPL = 0 (i.e., did not apply).

Procedures for assigning values to the Mix Score, Apply Score, Repair Score, and PPE Factor are described below.

10.2 Mix Scores

If applicators personally mixed pesticides more than 50% of the time (APCTMIX = 3), then they receive a mix score of 50. If applicators mixed less than 50% of the time (APCTMIX = 2), then their mix score is set to 20. Finally, if applicators reported never mixing (APCTMIX = 1), then they are given a mix score of 0.

If the response to the question on frequency of mixing is missing, the mix score is set to missing, except for applicators who responded to the questions about use of additives when mixing on the Take-Home Questionnaire. If the applicator indicated he or she used an additive when mixing either herbicides, insecticides for crops, insecticides to animals or fungicides, the mix score is set to 50. If the applicator indicates he or she did not mix for all four of these application types, the mix score is set to 0. Applicators with missing information on mixing are assigned a mix score of 20 when calculating their algorithm scores.

10.3 Apply Scores

In the Phase I questionnaire applicators reported application methods by checking a box by all of those methods that they used. Some application methods are not applicable to certain classes of pesticides. Thus, a separate application scores is assigned for the four different classes of pesticides and the score should only take into account those application methods that are relevant, as outlined in the paper by Dosemeci et al. 2002. For example, for herbicides, the relevant application methods to be considered are aerial application, distribution of tablets, banded application, broadcast application, hand spray, and backpack spray. Table 10-1 lists the application methods to be considered according to the class of pesticide used. If several relevant application methods are reported, the application score is obtained by averaging the weights for each method. The weights assigned to each method are given in Table 10-2.

For insecticides, two application scores are calculated: one for application of insecticides to crops and one for application of insecticides to animals. For calculation of intensity-weighted days for insecticides, the crop intensity score is used, except for the following insecticides: Coumaphos, Dichlorvos and Permethrin applied to animals. These insecticides are only applied to animals, so the insecticide to animal intensity score is used for calculation of intensity weighted days. For Chlorpyrifos, the insecticide to animal intensity score is used for applicators who did not report raising any crops, the crop intensity score is used for applicators who did not report raising any animals, and the average of the crop and animal intensity score is used for applicators who reported raising both crops and animals in Q31.

Table 10-1. Application methods to be considered for application score according to pesticide class.

Pesticide class	Relevant application methods
Herbicides	aerial, distributed tablets, banded/in furrow, broadcast/boom, hand spray, backpack
Crop insecticides	aerial, application to seeds, distributed tablets, banded/in furrow, broadcast/boom, hand spray, backpack, airblast, mist blower/fogger
Animal insecticides	ear tags, inject animal, dip animal, spray animal, dust/pour on animal, powder duster
Fungicides	application to seeds, distributed tablets, banded/in furrow, broadcast/boom, hand spray, backpack, airblast, mist blower/fogger
Fumigants	gas canister, row fumigation, pour fumigant

Table 10-2. Application scores given to each method and variables indicating use of the method.

Method	Variable name	Application score
Aerial	aapmth7	10
Apply to seeds	aapmth9	10
Ear tags	aapmth18	10
Distribute tablets	aapmth10	10
Banded/In furrow	aapmth8	20
Inject animal	aapmth15	20
Gas canister	aapmth12	20
Boom/broadcast	aapmth3	40
Row fumigation	aapmth13	40
Dip animal	aapmth16	70
Spray animal	aapmth17	60
Hand spray	aapmth4	70
Dust/pour on animal	aapmth19	70
Backpack	aapmth5	80
Mist blower/fogger	aapmth6	80
Airblast	aapmth2	140
Powder duster	aapmth14	90
Pour fumigant	aapmth11	90

The information in Table 1 and Table 2 is combined and expanded in Section 10.7 by adding the variable names from the take home questionnaires. If none of the application methods used for herbicides (indicated by a "Y" in the herbicide column in the method row) were used by applicator, that applicator's application method score for herbicides is zero. Similarly, if none of the methods for crop insecticides, animal insecticides, fungicides, or fumigants were used by an applicator, then the application score for that method is set to zero.

Similar weights are assigned to the corresponding variables from the Private Applicator Take Home Questionnaire (the Farmer Questionnaire) and the Commercial Applicator Take Home

Questionnaire. Combining the information in Tables 10-1 and 10-2 and expanding them to include the take home questionnaire variables, results in the chart shown in Table 10-3. Note that if a method is indicated as being used in either the Enrollment Questionnaire or a take home questionnaire or both, it is counted only once.

10.4 PPE Factor

Question E17 on the Enrollment Questionnaire asked applicators to mark the types of PPE they typically used when mixing/loading or applying pesticides. The responses to this question are used to calculate a PPE reduction factor to reduce the exposure intensity scores based on the type of PPE reported during Phase I. The reduction for PPE ranges between 0% for no use of PPE up to 90% for use of chemically resistant or rubber gloves and at least three additional types of the PPE listed on the Enrollment Questionnaire. The reduction factor for applicators who reported use of chemically resistant gloves (denoted by the indicator variable APROTEQ7) is 60%. An additional 10% reduction is assigned for each additional item of PPE (denoted by the indicator variables APROTEQ3, APROTEQ4, APROTEQ5, and APROTEQ8) up to a limit of 3 additional items. No reduction is provided for use of leather/fabric gloves (APROTEQ6).

APROTEQ1	E17: Do not personally handle pesticides
APROTEQ2	E17: Never use protective equipment
APROTEQ3	E17: Cartridge respirator/gas mask
APROTEQ4	E17: Face shields/goggles
APROTEQ5	E17: Disposable clothing
APROTEQ6	E17: Fabric/leather gloves
APROTEQ7	E17: Chemical resistant gloves [glove]
APROTEQ8	E17: Other protective clothing

The PPE factor is calculated as follows.

If all of APROTEQ1 = 1 and APROTEQ2 through APROTEQ8 are all zero set PPE factor to 0.

If APROTEQ2 = 1 and APROTEQ3, APROTEQ4, APROTEQ5, APROTEQ7, and APROTEQ8 are all equal 0, then PPE factor = 1.

If all of APROTEQ1 through APROTEQ8 are missing then set PPE factor to missing.

Otherwise, PPE factor = 1 - (0.6 * APROTEQ7 +
0.1 * MINIMUM((APROTEQ3 + APROTEQ4 + APROTEQ5 + APROTEQ8), 3)

NOTE: Because of the construction of the questionnaire, it is not possible for any of APROTEQ1 through APROTEQ8 to have missing values without all of these variables having missing values.

10.5 Repair Scores

The repair score is assigned a value of 20 if applicators reported performing repairs themselves during the past year (ADOSLF5=1). Otherwise, they are assigned a score of 0 (ADOSLF5=0). Note that no applicators have missing values for this variable.

10.6 Phase I Lifetime Intensity-Days

To obtain intensity-weighted exposure-days to a particular pesticide, lifetime days of use of each pesticide is calculated as the product of the years of use time day per year of use, and multiplied by intensity score appropriate for the pesticide class.

10.7 Method Table for Application Scores and Application Method Variable Names

Method Name	Variable Name	Application Score	Herbicide	Crop Insecticide	Animal Insecticide	Fungicide	Fumigant
Aerial	aapmth7	10	Y	Y		Y	
	ahrbmth6	10	Y				
	ainsmth6	10		Y			
	afungmt6	10				Y	
Airblast	aapmth2	150	Y	Y		Y	
	ahrbmth2	150	Y				
	ainsmth2	150		Y			
	afungmt2	150				Y	
Apply to seeds	aapmth9	10		Y		Y	
Backpack	aapmth5	90	Y	Y		Y	
	ahrbmth5	90	Y				
	ainsmth5	90		Y			
	afungmt5	90				Y	
Banded/In furrow	aapmth8	20	Y	Y		Y	
	ahrbmth7	20	Y				
	ainsmth7	20		Y			
Boom/broadcast	aapmth3	40	Y	Y		Y	
	ahrbmth3	40	Y				
	ainsmth3	40		Y			
	afungmt3	40				Y	
Dip animal	aapmth16	50			Y		
	acaimt6	50			Y		
Distribute tablets	aapmth10	10	Y	Y		Y	

Method Name	Variable Name	Application Score	Herbicide	Crop Insecticide	Animal Insecticide	Fungicide	Fumigant
Dust/pour on animal	aapmth19	70			Y		
	acaimt4	70			Y		
	acaimt10	70			Y		
Ear tags	aapmth18	10			Y		
	acaimt2	10			Y		
Gas canister	aapmth12	20					Y
	afumgmt2	20					Y
Hand spray	aapmth4	90	Y	Y		Y	
	ahrbmth4	90	Y				
	ainsmth4	90		Y			
	afungmt4	90				Y	
Inject animal	aapmth15	20			Y		
Mist blower/fogger	aapmth6	90	Y	Y		Y	
	ahrbmth8	90	Y				
	ainsmth8	90		Y			
	afungmt8	90				Y	
Pour fumigant	aapmth11	90					Y
Powder duster	aapmth14	90			Y		
Row fumigation	aapmth13	40					Y
	afumgmt3	40					Y
Spray animal	aapmth17	70			Y		
	acaimt7	70			Y		