EXHIBIT 10

The Federal Government's Agricultural Health Study: A Critical Review with Suggested Improvements

George M. Gray , Bernard D. Goldstein (Chair) , John Bailar , Devra Lee Davis , Elizabeth Delzell , Frank Dost , Raymond S. Greenberg , Maureen Hatch , Ernest Hodgson , Michel A. Ibrahim , James Lamb , Terry Lavy , Jack Mandel , Richard Monson , Mark Robson , Roy Shore & John D. Graham

The Federal Government's Agricultural Health Study: A Critical Review with Suggested Improvements

George M. Gray,[1],* Bernard D. Goldstein (Chair),[2] John Bailar,[3] Devra Lee Davis,[4] Elizabeth Delzell,[5] Frank Dost,[6] Raymond S. Greenberg,[7] Maureen Hatch,[8] Ernest Hodgson,[9] Michel A. Ibrahim,[10] James Lamb,[11] Terry Lavy,[12] Jack Mandel,[13] Richard Monson,[14] Mark Robson,[2] Roy Shore,[15] and John D. Graham[1]

[1]Center for Risk Analysis, Harvard School of Public Health; [2]Environmental and Occupational Health Sciences Institute, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School and Rutgers, The State University of New Jersey; [3]Department of Health Studies, University of Chicago; [4]World Resources Institute; [5]Department of Epidemiology, University of Alabama at Birmingham; [6]Professor Emeritus, Oregon State University; [7]Medical University of South Carolina; [8]Department of Community Medicine, Mount Sinai School of Medicine; [9]Department of Toxicology, North Carolina State University; [10]Department of Epidemiology, University of North Carolina; [11]Jellinek, Schwartz & Connolly, Inc.; [12]Department of Agronomy, University of Arkansas; [13]Division of Environmental and Occupational Health, University of Minnesota School of Public Health; [14]Department of Epidemiology, Harvard School of Public Health; [15]Department of Environmental Medicine, New York University

ABSTRACT

The Agricultural Health Study (AHS) has approximately 90,000 pesticide applicators and their spouses enrolled in a number of studies to determine whether exposures to specific pesticides are associated with various cancers and other adverse health outcomes. Although the AHS was intended to be an integrated program of studies, some significant difficulties have emerged. In this report, we examine the design of the AHS, identify important program strengths and flaws, suggest various improvements in the program, and recommend ancillary studies that could be undertaken to strengthen the AHS.

Overall, the AHS is collecting a large amount of information on potential determinants of health status among farmers and farm families. A promising feature of

*   Corresponding author: Harvard Center for Risk Analysis, Harvard School of Public Health, 718 Huntington Ave., Boston, MA 02115; Tel: (617) 432–4341; Fax: (617) 432–0190; E-mail: ggray@hsph.harvard.edu.

Gray *et al.*

the AHS is the prospective cohort study of cancers among farmers in which the research design determines exposures prior to the diagnosis of disease. More effort needs to be devoted to reducing selection bias and information bias. Success of the cohort study will depend in part on follow-up surveys of the cohort to determine how exposures and disease states change as the cohort ages. The cross-sectional and case-control studies planned in the AHS are less promising because they will be subject to some of the same criticisms, such as potentially biased and imprecise exposure assessment, that have characterized the existing literature in this field.

Important limitations of the AHS include low and variable rates of subject response to administered surveys, concerns about the validity of some self-reported non-cancer health outcomes, limited understanding of the reliability and validity of self-reporting of chemical use, an insufficient program of biological monitoring to validate the exposure surrogates employed in the AHS questionnaires, possible confounding by unmeasured, nonchemical risk factors for disease, and the absence of detailed plans for data analysis and interpretation that include explicit, *a priori* hypotheses. Although the AHS is already well underway, most of these limitations can be addressed by the investigators if adequate resources are made available. If these limitations are not addressed, the large amounts of data generated in the AHS will be difficult to interpret. If the exposure and health data can be validated, the scientific value of the AHS should be substantial and enduring.

A variety of research recommendations are made to strengthen the AHS. They include reliability and validity studies of farmer reporting of chemical use, biological monitoring studies of farmers and members of farm families, and validity studies of positive and negative self-reports of disease status. Both industry and government should consider expanded research programs to strengthen the AHS.

Key Words: epidemiology, pesticides, farmworkers, health effects

INTRODUCTION

The Agricultural Health Study (AHS) was launched in 1993 by scientists at the National Cancer Institute, the National Institute of Environmental Health Sciences, and the Environmental Protection Agency. The primary impetus for the study is a concern that exposures to chemicals on the farm, particularly certain fungicides, insecticides, and herbicides, may be responsible for a variety of adverse health effects, including cancer, neurological damage, reproductive problems, immuno-logic defects, nonmalignant respiratory disease, kidney disease, and impairments to the growth and development of children (Alavanja *et al.,* 1996).

As a result of this concern, just over 90,000 farmers, commercial applicators of farm chemicals, and their families in two states, Iowa and North Carolina, have been enrolled in a long-term health study. Most of the data in the study are being obtained from farmers through self-administered questionnaires and telephone interviews. Numerous questions were already asked of enrollees regarding their experiences as a farmer, their patterns of chemical use, their lifestyles, and their current health status. For some diseases, such as cancer, some of the future health information about enrollees will be obtained from state-wide registries.

Federal Government's Agricultural Health Study

The AHS is not a single study. Although the population of primary interest is the farmers ("private applicators"), there will also be studies of the health of commercial applicators and the spouses and children of private applicators. The AHS includes studies with at least four different designs and makes use of a variety of data sources.

First, the main prospective cohort study is expected to follow the 90,000 enrollees for many years or until death, to determine whether use of particular chemicals or other features of the farm environment and personal behavior are associated with poor health outcomes. This main study will not be completed until sufficient numbers of the cancers of interest have occurred or, ultimately, until most of the enrollees have died and the collected data on health outcomes have been fully analyzed. Interim reports on the cohort can be expected when the frequency of specific health problems supports a quantitative analysis of the factors associated with these health outcomes. An important design feature of the main cohort study is that much of the information on chemical use is obtained from farmers via survey methods *prior to the diagnosis of disease*. Although some enrollees had chronic diseases when they entered the study, the AHS investigators should consider analyzing the data with and without inclusion of these prevalent cases of disease.

Second, cross-sectional studies are being undertaken to determine the prevalence of certain noncancer health outcomes among farmers and farm families. The three initial cross-sectional studies are investigating (1) history of spontaneous abortion, menstrual function, and fertility in young women; (2) menopausal states, reproductive history, and selected chronic diseases in older women; and (3) neurologic symptoms and visual impairment in farmer-applicators. A cross-sectional design entails comparing the prevalence of reported adverse health outcomes with the reported use of or exposure to specific chemicals. Telephone interviews of subsamples of the cohort are being used to compare those people who responded to take-home questionnaires and those who did not as well as to obtain the information to augment the cross-sectional studies of non-cancer health outcomes (Sandler, 1998).

Third, nested case-control studies are planned for a variety of diseases including non-Hodgkin's lymphoma, leukemia, and cancers of the prostate, brain, ovary, breast, lung, colon, and stomach (Agricultural Health Study, 1993). Farmers in the cohort who develop a particular disease will be compared with controls selected from the cohort. Unlike the main cohort study, the nested case-control studies may entail obtaining some information from farmers or next of kin after a disease has been diagnosed. The investigators will examine whether cases report greater use of agricultural chemicals than selected controls. Cases and controls may also be invited to complete more detailed questionnaires aimed at obtaining a better understanding of possible differences in their exposure to a variety of farm and nonfarm factors.

Finally, some effort is being undertaken to determine how much farmers and their families have been exposed to selected chemicals. Biological monitoring, which typically entails the collection and analysis of urine and/or blood samples for multiple chemicals, is expensive. Biomonitoring was originally proposed to take place at 200 farms. Pilot studies found low participation rates (about 23%) and higher costs than anticipated and thus the program of exposure assessment has been scaled back. The current experimental design calls for samples to be gathered

from people at 125 farms, but unexpected obstacles have surfaced in obtaining funds for even this reduced program of biological monitoring.

The design and implementation of any research program as large and complex as the AHS requires many tradeoffs and compromises. Not every analyst would make the same choices, but on the whole we commend the AHS investigators for making a variety of sound choices in the face of limited resources and a complex challenge. As we emphasize below, we are particularly enthusiastic about the prospective cohort study of cancer outcomes because it responds directly to some of the methodological weaknesses of prior epidemiologic studies of farmers and pesticides. Other aspects of the AHS, such as the cross-sectional studies of disease prevalence, have serious problems. In this report we focus on what the strengths and limitations of the various AHS studies are, how the AHS can be improved, and what steps can be taken by the government and industry to enhance what is being done in the AHS through complementary efforts.

Information about the AHS used in this review was obtained primarily from publicly available documents and information presented at the AHS's annual public Advisory Panel meetings.[1] We recognize that more detailed plans may have been made but are not publicly distributed. Although the cohort has already been defined and enrolled in the study, numerous decisions have yet to be made about how the data will be analyzed and how future surveys of the cohort will be refined and improved. Thus, the emphasis in our report is on two issues: those that can be addressed by the principal investigators of the AHS through expansions or modifications of the workplan and those that need to be understood as inherent limitations when the findings of studies are published and disseminated.

The report is organized as follows. Section 1 addresses "Data Sources, Response Rates and Data Quality". Sections 2 and 3 address "Pesticide Exposure" and "Pesticide Use", respectively. Section 4 examines "Risk Factors Other Than Pesticides". In Sections 5 and 6 we examine the "Study Design Issues" and "Data Analysis Plans". Section 7 summarizes our recommendations on how the study can be improved and what additional studies can be undertaken to advance the field.

DATA SOURCES, RESPONSE RATES, AND DATA QUALITY

The AHS includes four types of data that could play important roles in epidemiologic analyses: health outcome data, pesticide use and exposure data, and data on potential confounders (risk factors) for disease. In this section, possible limitations in the scope or quality of each type of data are identified, and we present some suggestions aimed at enhancing data quality. Since most of the data used in the study are based on surveys of farmers and members of farm families, we begin with a discussion of the response rates obtained for the AHS questionnaires (Tarone *et al.*, 1997).

Response Rates to Questionnaires

The target population for the AHS is all persons required by the states of Iowa and North Carolina to obtain a pesticide applicator license. This includes "private"

Federal Government's Agricultural Health Study

applicators (farmers) and "commercial" applicators. Both states require periodic re-training to maintain a license for either type of pesticide applicator. The enrollment questionnaire was given to all attendees at training courses in the two states over a 3-year period. A 3-year cycle for licenses assured that all users had a chance to enroll. In January, 1997, enrollment through training classes was completed.

Not all applicators at training sessions agreed to participate. Some special recruitment efforts were undertaken to increase participation rates. In Iowa, the response rate for the enrollment questionnaire was 81.9% for private applicators and 42.2% for commercial applicators. In North Carolina, 84.8% of private applicators enrolled and the study design did not include commercial applicators. Overall, enrollment questionnaire data are available from about 53,000 private applicators and 5,000 commercial applicators (out of about 76,000 possible). Questionnaire data have also been collected from about 32,000 spouses of farmers (about 73% of those eligible).[2]

After pesticide applicators filled out the enrollment questionnaire at the training session, they were given three supplemental questionnaires (applicator; spouse; female and family health) to complete at home and return. The AHS uses the supplemental questionnaires to enroll spouses and other family members. The response rates for the supplemental questionnaires are low. Overall, about 44% of *enrolled* applicators completed and returned the additional questionnaire (33.5% of all eligible applicators). The Spouse Questionnaire, or a telephone administered version, was completed by 73% of eligible spouses. The Female/Family Health questionnaire was returned by about 39% of female applicators or spouses of enrolled farmers (64.6% of enrolled spouses).

The questionnaires are the primary source of data for the AHS. The enrollment questionnaire, which is used to define membership in the cohort, gathers personal identifiers on the applicator and his or her spouse. It also asks about work on and off of the farm, frequency of use of 22 pesticide compounds (*e.g.,* ever/never used and frequency of application) and ever/never used information on 28 more, one question about application methods and another about protective equipment, whether a doctor has ever diagnosed any of 16 diseases, and several questions on some lifestyle activities (including smoking) and the specific crops or livestock raised on the farm. These data are available for all applicators in the cohort except when there are missing responses.

The supplemental questionnaires are intended to gather more detailed information from the applicator and his or her spouse about pesticide use, family history of cancer, personal history of infectious and chronic diseases, over-the-counter medicine use, and diet. The Spouse Questionnaire, for the wife or husband of the applicator, asks for information about pesticide use and farm activities, along with information about factors such as laundering and vacuuming and information about the home that might influence pesticide exposure. Information about dietary and cooking practices is also collected. A self-reported medical history elicited from each subject includes about 55 diseases or disease symptoms. The Female and Family Health questionnaire is intended for female applicators or female spouses of pesticide applicators. This questionnaire collects information about the woman's reproductive cycle, pregnancies, and children. Identifiers, birthweight, nursing history, and whether the child ever worked on a farm are recorded for each child.

Gray *et al.*

The low and variable response rates to the supplemental questionnaires seriously affect the quality of the AHS. Steps have been taken to increase response rates but the rate of non-response remains substantial. We encourage more efforts to increase the response rate, to reduce the potential for selection bias and increase statistical power. An evaluation of the potential for selection bias to influence risk estimates should be undertaken.

In the prospective cohort study, low response rates to questionnaires designed to obtain information on subject identifiers, exposures, and baseline disease status will clearly diminish statistical power and may create bias. The success of the cohort study also depends upon acceptable response rates to future follow-up surveys of the cohort. Periodic follow-up surveys are necessary to determine how exposures and disease states change as the cohort ages, thereby maintaining the prospective character of the study. If low response rates occur with the follow-up questionnaires, the potential for bias will increase, partly from misclassification of subjects (and person-years) with regard to chemical exposure and partly from residual confounding stemming from inaccurate measurement of risk factors other than pesticides. According to the AHS protocol (Agricultural Health Study, 1993), follow-up questionnaires will be administered every 5 years. Since no follow-up has yet been administered, response rates are unknown.

Selection bias should be reduced in the prospective cohort study if persons who already have the disease(s) of interest are identified and excluded from the cohort at the beginning. Identification of diseases diagnosed at the time of enrollment into the cohort may be done well for conditions, such as some cancers, that have an easily defined point of diagnosis but is more difficult for certain neurological conditions and for renal, respiratory, and cardiovascular diseases. For instance, bias will occur if persons who are at risk of cancer and are exposed are more likely to participate by returning questionnaires. There are plans for cohort studies of kidney, neurologic, respiratory and cardiovascular disease that might be biased by the erroneous inclusion of subjects with disease onset before enrollment, if the probability of study participation depends on exposure status. Furthermore, if response rates are low for questionnaires designed to obtain information on medical conditions occurring during the follow-up period, the likelihood of bias is high.

In cross-sectional and case-control studies, low response rates have most of the same potentially detrimental effects on precision and accuracy as mentioned above. In addition, poor response raises the likelihood that selection bias will occur because it is likely that participation will depend both on exposure status and on "disease" status in a manner that could bias estimates of prevalence ratios or odds ratios (*e.g.*, through underrepresentation of exposed persons without disease).

Health Outcomes

Accurate ascertainment of the presence or absence of disease among farmers and members of farm families is critical to the success of the AHS. Some of the diseases of interest in the study are relatively rare and only a small number of cases of these diseases can be expected. Thus, it is appropriate to consider the quality of the health-outcome data being collected in the AHS, looking at the potential for both false-positive and false-negative errors.

Federal Government's Agricultural Health Study

Cancer

The statewide cancer registries in Iowa and North Carolina will be used by the AHS investigators to determine which subjects develop various types of cancer. The Iowa registry is well established and is believed to provide accurate and reasonably complete data on incidence of cancer in the state. The North Carolina registry is newer but should provide data of sufficient accuracy.

Use of the cancer registries will be hampered if personal identifiers such as name, birth date, Social Security number, and gender are not available from both the cancer registries and the AHS cohort. Such identifiers are critical to linking subjects in the AHS cohort to registry records. In February 1997 it was reported that in Iowa the four identifiers mentioned above were available for 94.0% of commercial applicators, 86.3% of private applicators, and 53.4% of enrolled spouses of married private applicators. In North Carolina the four identifiers were available for 86.6% of private applicators and 76.5% of spouses. Linkage with registries may be acceptable with current identifiers but AHS investigators are making efforts to increase the completeness and quality of data needed for record linkage. There are other ways to determine whether enrollees have developed cancer, but they are generally more expensive.

Non-Cancer Health Outcomes

Mortality from kidney, neurologic, respiratory, cardiovascular, and other diseases can also be assessed through objective measures that do not entail self-reporting by subjects in the cohort. For example, mortality from specific causes can be monitored through periodic follow-up through the National Death Index and state and local vital statistics records. Yet even for data from objective sources, potential validity problems need to be identified and addressed.

In order to accelerate the opportunity to cover a wide range of non-cancer outcomes, the AHS relies on self-reporting of health states by farmers and members of farm families on both the enrollment and supplemental questionnaires. The self-reporting occurs either through return of written questionnaires or responses to telephone interviews. Telephone surveys of special subgroups of the cohort are being employed to reduce the potential for selection bias in the cross-sectional studies, but it is possible for a modest amount of selection bias to have a substantial effect on results. Diseases of particular interest to the AHS investigators include kidney disease, neurotoxicity and neurological disease, reproductive and developmental impairments, and immunologic effects. Several questions ask about possible acute toxicity episodes associated with pesticide use.

Section IV of the main enrollment questionnaire has two questions regarding health. Question #28 inquires whether "a doctor has ever told you that you had any of the following conditions": A list of 16 conditions is supplied (asthma, tuberculosis, other chronic lung disease, pneumonia, melanoma of skin, other skin cancer, leukemia, Hodgkin's disease, non-Hodgkin's lymphoma, other cancer, heart disease, diabetes, Parkinson's disease, kidney disease, nervous disorder, and depression), each to be answered yes or no. For the cancer outcomes, it will ultimately be

feasible to compare the self-reports of subjects to the data obtained through the statewide cancer registries. A strategy for addressing discordant data is needed.

Question #29 inquires whether "your parents, brothers, sisters, or children related to you by blood ever had any of the following?" A list of 14 conditions is supplied, again with yes or no responses. If the subject has multiple blood relatives, the implicit understanding is that the question refers to any of them.

In addition to these questions about diseases, questions #14 and #15 provide additional information about acute health effects that may be related to pesticides. These questions do not ask about a medical diagnosis, and no effort is being made to validate the answers.

Question #14 asks "How often, if ever, have you had the following symptoms that you think may be related to your using pesticides?" There are seven listed symptoms: "been excessively tired", "had headaches/dizziness", "had nausea or vomiting", "had skin irritation", "had eye irritation", "had chest discomfort", and "felt nervous or depressed". For each symptom, the respondent is asked to respond on a scale of never/rarely, sometimes, frequently, almost always. This set of questions seems to combine elements of symptom frequency and causal attribution. It is not clear how the respondent is expected to judge whether such symptoms were "related to your using pesticides" unless the effects were immediate and unambiguous. It may be preferable to ask separate questions about the frequency of these symptoms and the respondent's view about whether they are associated with pesticide use, although questions about validity might remain. It is also not clear what would be meant by "frequently/almost always", since no frequency context is suggested to the respondent. The response may represent symptom frequency in absolute terms or as a percentage of the total number of pesticide applications. Given the ambiguous nature of this question, the meaning of the information that is elicited will be uncertain.

Question #15 asks subjects: "As a result of USING PESTICIDES (emphasis in original), how often have you: a. seen a doctor, b. been hospitalized." The possible responses are never, once, twice, or three or more times. Again, this question presumes that the respondent knows something about the causative role of pesticides in particular situations, perhaps because he or she experiences unusual symptoms in short order after the chemical is applied. Some visits may be after exposure but before symptoms appear. In ambiguous situations involving common symptoms and longer time lags, the respondent may not realize that the chemical exposure was responsible for the symptom or may attribute to the chemical a symptom that was not caused by the exposure. If the question is intended to provide a surrogate measure of exposure to chemicals, it needs to be used with caution if it is used at all.

Although there is limited information on noncancer health outcomes in the main enrollment questionnaire, the supplemental questionnaire includes a fairly detailed self-reported "medical history" from each subject. The low overall rate of response to the supplemental questionnaire, despite efforts to increase response, will prevent full understanding of the cohort's exposure and health states.

Questions #87 through #102 in the supplemental questionnaire ask about numerous aspects of the applicator's health status. For example, Question #87 asks about

Federal Government's Agricultural Health Study

each of 41 listed diagnoses (*i.e.*, whether a doctor has ever told the subject that he or she has that condition). Question #90 asks how frequently, during the last 12 months, the subject has experienced each in a list of 23 symptoms ranging from dizziness and headaches to feeling tense or depressed. Questions #96 through #102 focus on the respondent's vision and use of eyeglasses. Responses to these questions need to be validated.

It may be that any biases will cancel out because potential cases and non-cases interpret questions in roughly the same manner, as may be expected in a prospective cohort study, but it will be very difficult to know for sure the overall or net impact of any resulting biases.

Applicators who are women, and the spouses of male applicators, are also asked to complete a "Female and Family Health Questionnaire" that includes numerous questions on the subject's reproductive and pregnancy history, and about the health status of children. The AHS is also using a specialized "Women's Health Questionnaire" and a separate "Young Women's Health Questionnaire" to obtain specific pesticide use information and more detailed health information on subgroups of women who have enrolled in AHS. The former questionnaire has a special section on menopause while the latter questionnaire emphasizes menstrual functioning and pregnancy history.

Epidemiologists do not expect perfect concordance between self-reports and medical records. Although subjects may supply inaccurate data, medical records are themselves not free from error. The accuracy of self-reports presumably vary by type of health endpoint, questionnaire design, period of recall, and population studied. For many reproductive endpoints, the results of reliability and validity studies are reassuring, while for others there is concern (Bean *et al.*, 1979; Wilcox and Horney, 1984; Olson *et al.*, 1997). For some endpoints, such as menstrual function, there is no practicable gold-standard to compare with self-reports. It is important for the investigators to address how they will incorporate uncertainty about self-reports into their analyses and interpretation of results.

The AHS is collecting a large amount of self-reported health information on non-cancer health outcomes. Most of the specific questions on non-cancer health outcomes used in the questionnaires have not been assessed for validity or reliability and there appear to be no plans to initiate such studies by the AHS team. Apparently, follow-up questionnaires will not repeat questions about past health outcomes, preventing assessment of reliability. Some of these questions have already been used in previous studies and may have been subjected to some reliability and validity checks but study context can influence responses. More such studies would help users understand the quality of the non-cancer outcome information that will be analyzed in the AHS. It is crucial that reports of both the presence and absence of specific outcomes be validated in order to ascertain false-positive and false-negative errors.

Bias can occur when subjects know the purpose of a study and when they also know their exposure status and disease status. For example, "exposed" subjects (*e.g.*, heavy users of chemicals) with disease may be more willing to participate in the AHS cross-sectional studies than nonexposed subjects who also have the same disease. The prospective cohort design provides an important protection against such bias,

as long as a subject's exposure truly precedes the onset or diagnosis of the disease being investigated. The prospective cohort study also provides a good opportunity to obtain valid information on exposure and disease status. This strength may be enhanced through various analyses designed to detect and diminish information bias and other problems with the AHS data. It would probably be necessary to gather additional data to support such methodological substudies. It is preferable to minimize the potential for bias by obtaining valid information from all subjects.

PESTICIDE EXPOSURE

Although the primary goal of the AHS is to assess the relationship between human exposure to pesticides and a variety of adverse health outcomes, direct measurement of human exposure to pesticides will be limited by cost considerations. Most of the analyses will be based on surrogates for exposure.

As of early 1998, the U.S. USEPA team planned to select a sample of 125 farms and evaluate total exposure for several chemicals through measurement of environmental media, personal exposure (*e.g.,* through patches on clothing), and samples of urine and blood, taken soon after application. These samples will then be analyzed for a limited number of chemicals of greatest interest. Unexpected funding problems may prevent implementation of USEPA's plan.

Since no direct measures of pesticide exposure will be available on most of the 90,000 members of the AHS cohort, the investigators will rely primarily or exclusively on surrogates for pesticide exposure derived from the questionnaires administered to farmers and members of farm families. For example, previous studies have considered as surrogate factors such measures as frequency of application per year, number of years of application, and application practices that may be related to exposure (*e.g.,* method of application and type of protective equipment used) (Hoar *et al.*, 1986; Zahm *et al.*, 1990). It is not known how well any of these surrogates indicate biologically significant exposures or whether any is appropriate. A case can be made that exposure surrogates should be validated before initiating a major epidemiologic study, or at least before exposure-response analyses are undertaken.

A key goal of the USEPA portion of the AHS is calibration of reported work practices with actual farmer exposures, using the information obtained from the measurements gathered on the sample of farms. Ideally, this information would allow at least a ranking of exposure potential by method of application and protective equipment used. For example, some pesticides are formulated as liquids, and gloves may provide a great deal of protection. Others are formulated as dusts or sprays and thus gloves may make little difference, while a respirator or mask may greatly reduce exposure. Still others are large granules and neither type of protective equipment may have much influence on exposure.

Because of its limited size, the USEPA study is unlikely to provide a rigorous validation of the numerous exposure surrogates derived from the AHS questionnaire data. A larger sample of farms, pesticides, and work practices would be useful in validating the surrogates against the background of other significant determinants of exposure such as the subject's age and role in pesticide use. There are also

questions about the representativeness of sampled farms. The USEPA has had difficulty obtaining the participation of farmers. In a pilot study in North Carolina, fewer than 10% of farmers asked agreed to participate (U.S. Environmental Protection Agency, 1997). It seems unlikely that the farmers who agree to participate will be representative of all of the farmers in the AHS. The timing of the USEPA exposure study is also a source of concern. USEPA's exposure study is just getting underway but the AHS enrollment questionnaires have already been administered to the 90,000 enrollees. If the USEPA study raises questions about the validity of the exposure surrogates contained in the enrollment questionnaire, the data that have already been collected from farmers on work practices will be of diminished utility. If done in a timely fashion, it may be feasible to revise future follow-up questionnaires in a way that will benefit from the insights generated from the USEPA's exposure study.

Previous studies have relied on the assumption that total lifetime exposure to one or more pesticides is determined by annual frequency of application and number of years of application. Although this assumption may seem logical, there is no plan to validate it. It is possible that those farmers who apply pesticides frequently and have done so for many years do so with particular experience and care, which might suggest that their absorbed dose per application is less than the exposure of farmers who apply chemicals less frequently or have fewer years of experience in farming. Of course, bias may also run in the opposite direction if some applicators become careless or even contemptuous of risks as the substances and application practices become familiar. A particular task, such as mixing, may lead to much greater exposure than frequent application. If rare but serious mishaps or spills have a powerful influence on total lifetime exposure, number of applications may be a poor surrogate for total exposure, since the probability of mishap/spill may be smaller among high-frequency applicators. The USEPA study may not be large enough to detect these rare yet serious incidents. Thus, it is not obvious that total exposure to pesticides in a farmer's lifetime, on average, will be a straightforward multiple of the number of applications in a farmer's lifetime.

The use of inappropriate or imperfect exposure surrogates may compromise the validity of the study by producing erroneous measures of association. Errors due to misclassification can produce bias toward the null (attenuation of the magnitude of a true positive or inverse association) or bias away from the null (exaggeration of the strength of a true weak or true null association). In large prospective follow-up studies of relatively common exposures and diseases, exposure misclassification tends to be nondifferential with regard to disease status. Nondifferential exposure misclassification will produce bias toward the null if exposure is classified dichotomously (*e.g.,* exposed vs. unexposed, high vs. low exposure). If more than two categories of exposure are evaluated, however, nondifferential misclassification has an unpredictable impact and can produce bias away from the null (Correa-Villasenor, A., Stewart, W. F., Franco-Marina, F., and Seacat H. (1995); Thomas, 1995). In small studies or studies in which exposure is rare or disease rates are low, the impact of misclassification, again, is unpredictable. There is no guarantee that exposure misclassification will be nondifferential even if objective exposure assessment pro-

cedures are used. Misclassification will reduce the power of the study to detect any genuine cause-effect relationships and will also reduce the validity of findings. Reductions in power are a serious issue because they will undermine the ability of government and industry to regulate harmful exposures and to reassure farmers with "negative" results.

Biomonitoring studies of farmers who mix and apply pesticides with different frequencies might help resolve this matter, but such studies would need to be large in size and would be logistically complex. Such studies may induce behavioral changes (*e.g.*, extra safety precautions) on the part of some farmers that are not typical of their normal behavior.

Although it will be difficult to validate whether number of applications is a strong predictor of total exposure, it may be more feasible to study the impact of work practices and method of application on the amount of actual pesticide exposure. A farmer's personal habits can have an enormous influence on pesticide dose, as measured by urinary excretion, even when the same protective equipment is used (Lavy, Walstad, Flynn, and Mattice, 1982; Forbess *et al.,* 1982; Leng, Ramsey, Braun, and Lavy, 1982). It will be difficult to characterize this source of variability in the small sample of farmers being evaluated by the USEPA. Broader studies of the type planned by the USEPA, with a focus on the AHS pesticides and work practice and protective equipment questions, would be very useful. Some information on the role of work practices and protective equipment is already available in USEPA's Pesticide Use Handlers Database and our understanding is that the AHS investigators have begun to exploit this source of data. We encourage more efforts in this direction. The Department of Defense has conducted large programs of research on the efficiency, safety, and comfort of protective gear, and some of the results (*e.g.*, points of leakage or tolerance by the protected person) may be directly applicable to pesticide applicators.

There are also practical and technical concerns associated with any urine biomonitoring program. The USEPA investigators are aware of many potential pitfalls but still may have difficulty dealing with them. One of the biggest problems is time. If a pesticide is rapidly excreted, measurements must be made quickly after a single application to be useful for exposure assessment. If, however, the material is cleared slowly from the body, the amount of the chemical measured in urine will be highly dependent on the frequency of applications and the time interval between applications. There are significant differences in pharmacokinetics across compounds that will influence the relationship between frequency/pattern of use and exposure. Thus, a serious biomonitoring program must have a protocol that tailors the measurement regime to the behavior of the compounds under study. Yet the USEPA plans to sample only a fraction (perhaps as few as 10) of the 50+ chemicals being assessed in the AHS, and funding obstacles are jeopardizing even this modest level of effort.

Another key assumption of the AHS is that exposure of farm family members to pesticides is associated with the farmer's patterns and frequencies of use. Little is known about the nature of this relationship or how it varies for different compounds and farm types (Lowenherz *et al.*, 1997). The existing studies are small in size and are quite limited in the number and type of pesticides evaluated. Assuming partici-

pation obstacles can be overcome, biomonitoring could be used productively to better understand the presence and magnitude of indirect exposures to farm families that are assumed in the Spouse Questionnaire and the Female and Family Health Questionnaire. USEPA has limited plans in this area that will need to be expanded considerably if they are to be useful in the AHS.

The NCI also plans a biomarker component, collecting buccal DNA samples from a subsample of the AHS cohort, to store for later analysis of genetic polymorphisms potentially related to susceptibility to pesticide-induced disease. Although this effort is of considerable scientific interest, it is not likely to assist in validation of the exposure surrogates to be used in epidemiologic analyses.

In general, a major limitation of the current design of the AHS study is that so few direct measurements of human exposure to chemicals will be available. The information that USEPA plans to collect may be useful in its own right but, for the reasons stated above, is not likely to be as useful as it could be for use in the epidemiologic analyses to be performed in the AHS. Pesticide exposure studies that are linked to epidemiologic investigations are urgently needed if a major advance is to be made in our understanding of the relationship between pesticides and human disease. The significant cost associated with such an effort is noted, but the scientific value of this major epidemiologic study is questionable without a valid exposure assessment.

PESTICIDE USE

In the AHS, the questionnaires filled out by subjects elicit information on various aspects of pesticide use rather than on exposure directly. This approach is sensible because the respondent is in a better position to report accurate information on whether and how a chemical is used than information on the amount of exposure to chemicals. However, there are still serious questions about the quality of the pesticide use data that are being collected in the AHS. Since these data are likely to be critical to the interpretation of the epidemiologic analyses, the associated quality concerns need to be considered carefully.

In the AHS enrollment questionnaire, the primary questions (Qo, #11A-D) ask: "During your lifetime have you ever personally mixed or applied this pesticide?; how many years did you mix or apply this pesticide?; in an average year when you personally used this pesticide, how many days did you use it?; and when did you first personally use this pesticide?" (Paraphrased). These questions are posed for 22 named pesticides. For an additional 28 compounds, there is a simple question about whether that pesticide had ever been used.

In order to answer these questions, respondents must remember with some accuracy when they first used products and their frequency of use of each pesticide product, and they must be able to compute averages in their head involving multiple years of use. For older subjects who have many years of farm experience, accurate responses will be difficult to supply. Moreover, some pesticides are sold and applied as mixtures and thus the exact ingredients may not be known to farmers. It can reasonably be expected that there will be inaccuracies in these data.

Gray *et al.*

In the AHS enrollment questionnaire, there are two important questions about work practices. Question 16 asks: "how do you personally apply pesticides?" The offered answers include 20 options that are not differentiated by livestock or crop farming, by specific crop, or by pesticide used. Question 17 asks "what type of protective equipment do you generally wear when you personally handle pesticides?" The offered answers include 8 options, again making no distinction between farm types or pesticides used. Since most farmers will have had different practices for different crops or pesticide products, it is not clear how they will answer these questions in a meaningful way since multiple answers do not appear to be allowed.

There are, of course, real concerns about the ability of farmers to recall use of specific pesticides, let alone their frequency of use, when confronted with a long list of compounds. Many farmers know pesticides by trade names, not technical names. The AHS questionnaires list some trade names for all chemicals but the list is not exhaustive. In addition, farmers now often use formulations that contain several pesticides. A respondent who knows only one of the compounds or trade names could underreport the use of other pesticides in the mixture. Errors of recall may occur differentially between controls and diseased persons.

Due to a change in enrollment procedures, the AHS investigators do have duplicate enrollment questionnaires from 1223 applicators from Iowa (Alavanja, 1998). Reliability was reported as both the percent agreement (the fraction of applicators giving the identical answer to a question on both questionnaires) and kappa statistic, often used as a measure of reliability. For example, smoking had an agreement of about 90% and a kappa of 0.88. Reports of ever/never use of specific pesticides had agreement around 80% with kappas around 0.60. The agreement of frequency of use questions was not reported. Some questions, especially those about vegetable and fruit consumption, had quite low agreements (30 to 40%) and kappas (about 0.50). Of course, this analysis does not address the validity of the responses. It may be useful to include some more important use questions on future follow-up surveys to gauge reliability in the whole cohort.

A weakness of the AHS is that adequate information is not being collected on excipients such as solvents, stabilizers, diluting agents, preservatives and other chemical substances that are used with pesticide products. Confusion may occur about whether reported health effects are attributable to active ingredients or excipients. For regulators and firms interested in the design of pesticide products, it is crucial to know what precisely is causing a reported health effect.

There is no reason to believe that large numbers of subjects were deliberately dishonest in the enrollment questionnaire about their patterns of pesticide use. However, the questions about use of protective equipment may have induced some "socially desirable" but inaccurate answers, especially when questionnaires were administered at training sessions. It is also quite possible that pesticide products near the bottom of the lists of 22 and 28 were checked less frequently by respondents who became weary filling out this rather arduous aspect of the questionnaire. This problem could be smoothed out in the future follow-up surveys by rotating the order of the products.

A study of the magnitude of the AHS requires good understanding of the validity and reliability of each major data set. The AHS will obtain pesticide use data from

Federal Government's Agricultural Health Study

responses to a written questionnaire of farmers. Data will be collected both at the beginning of the study and with follow-up questionnaires of unspecified frequency, for either the whole cohort or a select subsample, in later years. On the subject of validity, purchase records have been used in the past to ascertain whether written answers to a "yes/no" question on use of specific products are accurate. One study reported a 60% agreement rate between purchase records and reported use of specific products (yes/no) (Blair and Zahm, 1993). Agreement between farmer's recall of years of use and the records of their suppliers ranged from 38% to 68% depending on type of pesticide and crop. Measures of frequency of use in a year have never been subjected to a validation study.

When social scientists find it difficult to validate questionnaire data, it is typical to at least conduct reliability studies, such as repeated administrations of the same (or similar) questionnaire(s) to respondents, to determine whether answers to the same question are stable. Few reliability studies of self-reported pesticide use, particularly the quantitative responses, have been published in the literature (Johnson *et al.*, 1993). In addition, since reliability is influenced by the particular wording of questions and response choices, there probably would be limited generalizability from reliability studies of other questionnaires.

The questions of reliability and validity regarding the reported data could be addressed in several ways. In addition to the small study already mentioned, a comparison of the responses of farmers to selected questions that have been included on both the enrollment and supplemental questionnaires will provide some ideas about reliability. Studies comparing self-reported use to purchase records for a subsample of the AHS farmers could provide an idea of the validity of self-reported use estimates. Even if recent purchase data can be obtained, it is likely that purchase records for earlier years will be less complete. Thus, it will be more difficult to verify the accuracy of self-reports of pesticide use in the past. Another opportunity to check self-reports might come from the Extension Service recommendations for each crop in Iowa and North Carolina. Consistency between self-reports and the recommendations of the Extension Service is one possible measure of accuracy. However, if such recommendations are widely known, farmers may be reluctant to report actual use patterns that deviate significantly from these recommendations.

The chemicals, formulations, and application methods used on farms have changed significantly over time. Herbicides once applied at rates of pounds of active ingredient per acre are now applied in ounces per acre. Formulations have been developed to reduce exposure by making the pesticide in large granules or as packets that are dropped into an application tank, with no need for mixing or loading. These changes in patterns of pesticide use mean that data gathered about farming practice today are not a valid reflection of what was done in the past. The amount of exposure per application is probably smaller today than it was years ago, further complicating any calculation of cumulative exposure.

These details are important because if pesticides cause chronic diseases such as cancer and neurological disease, the biologically meaningful measure of exposure may be a cumulative dose figure that accounts for farming practices years or even decades ago. For chronic diseases diagnosed over the next 5 years or so, the exposure of interest probably occurred many years ago. Yet information about

Gray *et al.*

changes in farming practices over time is not being gathered in the AHS. In addition, the extent of pesticide use information to be collected in follow-up surveys of the cohort is not clear. If most of the pesticide use assessment in the AHS proves to be retrospective, the AHS will have little advantage over previous studies.

RISK FACTORS OTHER THAN PESTICIDES

Numerous factors other than pesticide use are known or suspected to contribute to the development of various diseases and health impairments under study in AHS. These factors are important because they may confound (exaggerate or attenuate) the effects of pesticides, they may interact with the effects of pesticides, or they may prove to be of much greater quantitative importance than pesticides even if they are not confounders or interacting variables.

Confounding Variables

In epidemiologic analysis, a confounding variable is a risk factor for the disease of interest that is associated with the exposure of interest (in this case, pesticides). For example, in an analysis to determine whether frequent application of a particular pesticide is a risk factor for a particular type of skin cancer, exposure to sunlight is a potentially confounding (or interacting) variable. The ultraviolet radiation from exposure to sunlight is known to be a cause of skin cancer and farmers who engage in frequent application of pesticides may have more exposure to the sun than other farmers. If exposure to sunlight is a confounding variable and is omitted from the epidemiologic analysis, the estimated risks associated with pesticide exposure will be biased. This bias can be reduced or eliminated by collecting information on the confounder and including such information in a multivariate analysis of the disease in question.

Concern about possible confounding may arise if certain patterns of pesticide misuse (*e.g.*, failure to use protective equipment) are used as a surrogate for pesticide exposure without consideration of the farmer's lifestyle. Farmers who do not use protective equipment (or engage in risky application practices) may be more likely to engage in a wide range of risky behaviors at work and at home than farmers who use protective equipment (or engage in low-risk application practices). Some of those risky personal actions may be linked to the health outcomes under study.

The AHS collects data on numerous variables that might confound the relationship between pesticide use/exposure and disease outcomes. Yet we know of no effort to identify such confounding variables and include them in the AHS study plans. Information about some risk factors other than pesticides is being collected in the AHS study (*e.g.*, aspects of the diets of farmers) but it is not clear whether such variables are correlated with pesticide exposure and are likely to cause the same types of tumors that chemicals may cause. In addition, since these data are collected in the supplemental questionnaires, they are not available for the entire cohort.

Interacting Variables

The effects of pesticide exposure on human health may be magnified or attenuated by other behavioral and/or environmental factors. For example, it has been

shown that the risk of lung cancer due to radon exposure among uranium miners is much larger among smokers than nonsmokers (Hornung, Deddens, and Roscoe, 1995). We do not know of any interaction effects to be expected in the AHS data, but if others know or suspect of such interactions, they should be postulated explicitly prior to data analysis and then tested rigorously in the statistical analyses.

Other Important Risk Factors

Although pesticide exposures are certainly worthy of study, these exposures are not necessarily the most biologically plausible determinants of disease in farmers or farm families and they may not prove to be as quantitatively important as a variety of risk factors such as smoking, diet, and obesity. Even accepting that chemicals are a major priority for study, more effort might be devoted to understanding farmer exposures to such agents as veterinary pharmaceuticals, engine oils, consumer products, animal viruses, and the crops themselves.

If modified appropriately, the AHS could be used to generate comparative information that might help farm families develop a sense of perspective about the relative risks associated with different risk factors in farm life. In order to serve this function, future surveys of the cohort planned by the AHS investigators need to devote more attention to risk factors other than pesticides and compare their relative significance to those of pesticides based on rigorous epidemiologic analysis. Nevertheless, a significant focus on pesticides is worthwhile.

STUDY DESIGN ISSUES

From a methodological perspective, the AHS employs several different study designs in various phases of the epidemiologic inquiry. They include a prospective cohort design, a nested case-control design, and a cross-sectional design. These different study designs have inherent strengths and weaknesses that need to be understood when the findings of the study are interpreted and compared to the findings of other investigators.

Prospective Cohort Study

A typical prospective cohort study follows subjects from the time of enrollment in a study until a particular disease is diagnosed or some other event occurs and/or death. The subjects' frequency and/or degree of exposure to the chemical or physical agents of interest are typically documented at the time of enrollment and throughout the follow-up period. An advantage of this study design is that exposure determinations are made by the investigators before anyone (including the investigators and the subjects) knows which subjects will develop a particular disease or die prematurely. A disadvantage of the prospective design is that accurate measurement of exposure to pesticides and other disease determinants requires that the cohort be questioned or monitored at intervals during the study period, not just at the beginning. For cancers diagnosed during the first 5 years of study, the exposure assessment in the cohort study is based on recollections of pesticide use patterns from years or even decades ago.

Gray *et al.*

Determining exposure status prior to knowledge of health outcome is particularly critical in the epidemiology of pesticides. Previous findings in the literature, which were based primarily on the case-control design, have been criticized on the grounds that those farmers who developed disease (or their next of kin) may have been motivated (for a variety of reasons) toward more complete and accurate reporting of pesticide use and/or exposure than those farmers who did not develop the diseases of interest (Ibrahim et al., 1991). If such differential misclassification of exposure occurs, it will tend to create a spurious positive association between exposure and disease. The prospective cohort design selected by the AHS investigators reduces, but does not eliminate, the chances that bias from differential exposure misclassification will occur because use and exposure are determined prior to knowledge of health outcome. It is critical that follow-up surveys of the cohort be administered on a regular basis to document how exposure and disease states change as subjects age.

The major disadvantage of the prospective cohort design is that, for some chronic illnesses, it takes a long time for sufficient numbers of subjects to fall ill or for the data to be useful for analysis. It is also an inefficient approach to studying relatively rare tumors such as soft-tissue sarcoma and leukemia. Overall, though, we are very enthusiastic about the decision of the AHS team to invest in the prospective cohort design and encourage the investigators to make every feasible effort to achieve acceptable response rates in the follow-up surveys of the cohort and address potential biases in the study.

Nested Case-Control Study

A typical case-control study will enroll "cases" who are known to have the disease in question and compare them to a random subset of "controls" who do not have the disease in question. If cases and controls are both selected from subjects enrolled in a particular cohort study, the study is referred to as a "nested" case-control study. The strength of this design is that the cases are included in the cohort studied. If exposures to a particular agent cause the disease in question, then the life histories of the cases should exhibit different (and presumably greater) exposures than the life histories of controls. Exposures to the agents of interest are typically assessed retrospectively for cases and controls (*i.e.*, after the death has occurred or the disease determination has been made), sometimes via interviews with next of kin or through reconstruction of job histories and practices. Like the prospective study, the nested case-control aspect of the AHS would be constrained by the time to development of disease and the numbers of persons in the cohort. We do not discuss this design in detail here, because it is currently being given low priority in the AHS and its strengths and weaknesses have been addressed elsewhere (*e.g.*, Monson 1990).

The Cross-Sectional Design

A typical cross-sectional study collects information on exposure and disease simultaneously from a sample of subjects. The association between reported exposure and disease is then investigated within the sample. If exposure causes disease,

Federal Government's Agricultural Health Study

it is anticipated that subjects that report more exposure will be more likely to report the health outcome of interest. A major weakness in the cross-sectional approach is the potential for bias. When exposure and disease data are gathered at the same time, it may also be unknown whether the temporality is correct, that is, that the exposure to the presumed cause actually preceded the disease, especially with diseases with no easily identified time of onset. There may be bias if persons change exposure status because of disease (*e.g.*, people with disease may stop working with pesticides). Although this research design has some utility for generating hypotheses, it is not considered useful in defining most cause-effect relationships. There are also concerns about the quality of the data gathered for the AHS cross-sectional studies. Without medical verification of self-reported disease, any associations found in the cross-sectional studies will be a weak basis for planning future studies.

DATA ANALYSIS PLANS

While the AHS study team presumably has some well-defined primary hypotheses, they are not specified in the *Environmental Health Perspectives* article (Alavanja et al., 1996) or in the more detailed study plans that have been made available to the public. By well-defined primary hypotheses, we refer to *a priori* hypotheses regarding specific chemicals, specific tumor types or health outcomes, and specific surrogate measures of exposure. Specific hypotheses and detailed plans for analysis help focus the gathering of data on both exposure and disease outcomes. They may also help investigators to avoid overinterpretation of the random oddities that occur in any large and complex data set.

Given the many possible comparisons of pesticides, methods of use, work practices, and health outcomes, a formal statement of why a particular pesticide/outcome combination should be analyzed seems desirable. Without any precommitment to specific hypotheses, the proper interpretation of any associations that are found will be less clear. Although it is appropriate for the AHS team to explore many possibilities when the data are analyzed, it should be clear to readers and decision makers which results confirm prior evidence or concerns and which are found only in the AHS data.

The large amount of questionnaire data developed by the AHS provides rich scientific opportunities but also particular challenges for analysis and interpretation. For example, information is gathered from respondents on numerous health outcomes (approximately 25 outcomes in the private applicator enrollment questionnaire, 70 outcomes in the farmer applicator and spouse questionnaires, and 35 outcomes in the female and family health questionnaire — a total of 130). For cancer, there will be numerous tumor types available for analysis from registry data. In addition to numerous health outcomes, information is gathered on numerous pesticide products (approximately 50 in the enrollment questionnaire and another 100 in the farmer applicator questionnaire). For exposure (dose)-response analysis, it appears that more than 35 different surrogates of exposure can be constructed from the responses to the questions about pesticide use, application methods and

Gray *et al.*

work practices (*e.g.*, average days of use per year, number of products used, years of use, different types of protective equipment and methods of application).

One can confidently predict that some of the multitude of exposure-response combinations will be statistically significant in the absence of any real effect. Without clearly stated *a priori* hypotheses, the investigators will have to exercise considerable discretion in data analysis and may exercise insufficient or excessive caution in their interpretation. The exercise of this discretion can be evaluated by the scientific community only if a small number of completely specified primary hypotheses are developed prior to any inspection of results. "Completely speci-fied" means that the method of analysis must be given in detail for each primary hypothesis. The benefit of this approach is the increase in plausibility of any "positive" findings among the primary hypotheses; the cost is that all other hypoth-eses lose some support, though some may still be compelling and others may be examined in subsequent studies.

Important questions arise about the role of conventional measures of statistical significance in the reporting and interpretation of results. Should numerical adjust-ments be made to published p-values to account for multiple comparisons? Given that many possible associations may be explored prior to publication of final results, what degree of documentation should be provided by the investigators of explor-atory analyses? If the documentation requirements are minimal, how will the scien-tific community understand the importance of the associations that are reported? The importance that may be placed on findings of no association between a specific pesticide and health outcomes raises the question of the reporting requirements for analyses that failed to find an association. Parallel consideration must be given to reporting requirements for "inverse" associations (*e.g.*, relative risks less than 1.0 for a particular exposure). At the same time, it would be helpful if the AHS investigators would publish all data and analytical results in some accessible format. Key findings would especially benefit from documentation of their consistency within the AHS database. Widely accessible electronic media such as the World Wide Web makes this feasible.

A detailed analysis plan and careful interpretation can reduce or eliminate these concerns. Examination of internal consistency can provide information about the plausibility of a particular association. A reasonably consistent dose-response gradi-ent is an important criterion. One implication of this criterion is that statistically significant dose-response trends caused primarily by one dose group, especially if it is an intermediate dose group, should be interpreted cautiously. On the exposure side, a finding that the strength of an association increases with particular use practices that are expected to yield higher exposures (and decreases with increasing farmer care), could be valuable evidence in buttressing study results. Sensitivity analyses involving different exposure surrogates and exposure groupings can also demonstrate whether findings are robust.

A key form of evidence to inform hypotheses and corroborate (or refute) analytic findings is biological plausibility. Pesticides, in addition to prescription drugs, are among the most thoroughly studied of all chemicals from a toxicologic perspective. Pesticides are diverse in mode of action and in excipients, raising doubts about attempts to group pesticides for analysis except under very specific conditions (*e.g.*,

examining insecticides with similar mechanisms of *toxic* action and with similar excipients). In the interpretation of epidemiologic results, dose considerations from toxicology can play an important role in determining the plausibility of the response. Associations with exposures far below those causing effects in animals may be less credible than those demonstrated at higher exposure levels. The nature and limited amount of exposure information in the AHS makes this important use of toxicology difficult. More generally, toxicological reasoning has not yet played a significant role in the design or execution of the AHS but should be an important part of a detailed plan of analysis.

SUMMARY OF RESEARCH RECOMMENDATIONS

The AHS is a major undertaking with the potential to add significantly to our knowledge of possible associations between pesticide use and other factors and the health of farmers. The weight that will be accorded to results from this major study requires care in assuring the accuracy of the findings. Several of the most important limitations of the AHS could be addressed through additional research with the cohort or through complementary studies on different groups. The priorities should be to (1) assess the validity of self-reported health outcomes; (2) explore the reliability and validity of pesticide use data; (3) understand the relationship between exposure surrogates and exposure; (4) examine the biological plausibility of any hypotheses; and (5) develop explicitness on analysis and statistical issues.

Assessing the Validity of Self-Reported Health Outcomes

Many of the early analyses from the AHS will be based on self-reported health data. The validity of these data is crucial to interpretation of the results. There are studies in the literature that raise serious questions about self-reports of disease (Harlow and Linet, 1989; Paganini-Hill and Chao, 1993; Kehoe et al., 1994; The Italian Longitudinal Study on Aging Working Group, 1997). Clinical verification of key self-reported health outcomes, where feasible, is essential. It is important that validity be assessed for both those members of the cohort reporting disease and those who claim none. These studies could also help address some concerns about recall bias in the noncancer studies as well as concerns about whether the disease was indeed preceded by exposure.

Exploring the Reliability and Validity of Pesticide Use Data

Since pesticide use data will be the basis for categorizing potential pesticide exposure in the AHS, the validity of these data is also critical. A simple and pertinent step would be to readminister the questionnaire to a sample of respondents to see how much the answers change. Other studies to validate reported pesticide use, for example, by comparison with purchase records, are also essential. A relatively simple check would consist of questions about number of acres for each specific crop for which a specific pesticide was used. This would allow comparison to label instruc-

tions or Extension Service recommendations to help gauge the validity of use reports. Results of validation studies would suggest the amount of confidence that we could place in the questionnaire data as well as pinpoint ways to enhance the design of follow-up questionnaires. Validation studies will be able to address only relatively recent use since use records from the past are likely to be less complete. Given that many of the pesticides of concern for cancer were used more heavily in the past, and that a substantial period occurs between exposure and detection of disease, there may be significant questions about the validity of self-reported pesticide use in earlier years.

Understanding the Relationship between Exposure Surrogates and Exposure

Complementary studies are needed to assess the accuracy of the assumptions in the AHS that link specific use patterns and work practices with different levels of exposure. Biomonitoring studies could provide critical information to link pesticide use information to actual exposure by measuring pesticide levels in the blood or urine. Biomonitoring studies to correlate farmer exposure and dose to pesticide use patterns and work practices would be extraordinarily valuable in linking chemical use data to exposure categories. Similarly, biomonitoring studies of spouses and children of farmers could help determine whether conditions of pesticide use are associated with family exposures that are frequent enough and high enough to lead to possible adverse effects. This effort would help focus attention and resources on the most critical of possible adverse effects.

Assessing the Biological Plausibility of Any Associations

A key research need is the careful enumeration, in advance of analysis, of the biological effects expected at relevant doses for specific pesticides. This undertaking will help avoid the criticism that identified associations are supported only by toxicologic explanations that are *post hoc* and hence unreliable. This effort should rely on both the existing epidemiologic literature and the immense toxicologic database on pesticide products. Dose-response information must play a key role. Identification of chemicals expected to be capable of affecting health at anticipated exposures can corroborate findings and help focus analysis efforts.

Analysis and Statistical Issues

It is critical that a detailed analysis plan for the AHS be developed. Specifics to be addressed should start with a small number of precise hypotheses about pesticide/disease relationships, including in detail the analytic method. Potential confounders, interacting variables, and other risk factors should be identified in a systematic way, where possible, with a focus on causation of specific diseases. There is a need to specify an analytic framework, including specific statistical procedures, that encompasses decision rules for analysis and reporting.

The general study plan of the AHS is not yet detailed enough to support a confident evaluation of the technical strengths and weaknesses of this major undertaking, and we recommend substantial efforts toward developing such a plan. The level of effort and detail we are suggesting here would be typical of a major

Federal Government's Agricultural Health Study

investigator-initiated proposal that is peer reviewed and judged to be worthy of funding by the National Institutes of Health.

ACKNOWLEDGMENTS

REFERENCES

Agricultural Health Study Protocol, Revision of December 16, 1993.

Alavanja, M. C. R. Agricultural Health Study Information Packet for Advisory Panel Meeting, March 4–5, 1998 — NCI Progress Report.

Alavanja, M. C. R., Sandler, D. P., McMaster, S. B., Zahm, S. H., McDonnell, C. J., Lynch, C. F., Pennybacker, M., Rothman, N., Dosemeci, M., Bond, A. E., and Blair, A. 1996. The agricultural health study. *Environ. Health Perspect.* **104**, 362–369.

Bean, J. A., Leeper, J. D., Wallace, R. B., Sherman, B. M., and Jagger, H. 1979. Variations in the reporting of menstrual histories. *Am. J. Epidemiol.* **109**, 181–185.

Blair, A. and Zahm, S. H. 1993. Patterns of pesticide use among farmers: implications for epidemiologic research. *Epidemiology* **4**, 55–62.

Correa-Villasenor, A., Stewart, W. F., Franco-Marina, F., and Seacat H. 1995. Bias from nondifferential misclassification in case-control studies with three exposure levels. *Epidemiology.* **6**, 276–81.

Forbess, R. C., Morris, J. R., Lavy, T. L., Talbert, R. E., and Flynn, R. R. 1982. Exposure measurements of applicators who mix and spray paraquat in grape vineyards. *Horticult. Sci.* 956–957.

Harlow, S. D. and Linet, M. S. 1989. Agreement between questionnaire data and medical records. *Am. J. Epidemiol.* **129**, 233–248.

Gray *et al.*

Hoar, S. K. *et al.* 1986. Agricultural herbicide use and risk of lymphoma and soft tissue sarcoma. *J. Am. Med. Assoc.* **256**, 1141–1147.

Hornung R. W., Deddens, J., and Roscoe R. 1995. Modifiers of exposure-response estimates for lung cancer among miners exposed to radon progeny. *Environ. Health Perspect.* **103**(Suppl. 2), 49–53.

Ibrahim, M. A. *et al.* 1991. Weight of the evidence on the human carcinogenicity of 2,4-D. *Environ. Health Perspect.* **96**, 213–222.

Johnson, R. A., Mandel, J. S., Gibson, R. W., Mandel, J. H., Bender, A. P., Gunderson, P. D. and Renier, C. M. 1993) Data on prior pesticide use collected from Self- and proxy respondents. *Epidemiology* **4**, 157–164.

Kehoe, R. *et al.* 1994. Comparing self-reported and physician-reported medical history. *Am. J. Epidemiol.* **139**, 813–818.

Lavy, T. L., Walstad, J. D., Flynn, R., and Mattice, J. D. 1982. (2,4-Dichlorophenoxy)acetic acid exposure received by aerial application crews during spray operations. *J. Agric. Food Chem.* **30**, 373–381.

Leng, M. L., Ramsey, J. C., Braun, W. H., and Lavy, T. L. 1982. Review of Studies with 2,4-Dichlorophenoxyacetic Acid in Humans Including Applicators Under Field Conditions. ACS Symposium Series No. 182. (Plimmer, J. R., Ed.), American Chemical Society.

Lowenherz, C., Fenske, R. A., Simcox, N. J., Bellamy, G., and Kalman, D. 1997. Biological monitoring of organophosphorous pesticide exposure among children of agricultural workers in central Washington state. *Environ. Health Perspect.* **105**, 1344–1353.

Monson, R. R. 1990. *Occupational Epidemiology,* 2nd ed. CRC Press, Boca Raton, FL.

Olson, J., Shu, X., Ross, J., Pendergrass, T., and Robison, L. 1997. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. *Am. J. Epidemiol.* **145**, 58–67.

Paganini-Hill, A. and Chao, A. 1993. Accuracy of recall of hip fracture, heart attack, and cancer: a comparison of postal survey data and medical records. *Am. J. Epidemiol.* **138**, 101–106.

Sandler, D. P. Agricultural Health Study Information Packet for Advisory Panel Meeting, March 4 to 5, 1998 — NIEHS Progress Report.

Tarone, R. E., Alavanja, M. C. R., Zahm, S. H., Lubin, J. H., Sandler, D. P., McMaster, S. B., Rothman, N., and 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. Indust. Med.* **31**, 233–242.

The Italian Longitudinal Study on Aging Working Group 1997. Prevalence of chronic diseases in older italians: comparing self-reported and clinical diagnoses. *Int. J. Epidemiol.* **26**, 995–1000.

Thomas D. C. 1995. re: When Will Nondifferential Misclassification of an Exposure Preserve the Direction of a Trend? *Am. J. Epidemiol.* **142**, 782–4.

USEPA (U.S. Environmental Protection Agency). Agricultural Health Study Information Packet for Advisory Panel Meeting, February 27 to 28, 1997 — USEPA Summary.

Wilcox, A. and Horney, L. 1984. Accuracy of spontaneous abortion recall. *Am. J. Epidemiol.* **120**, 727–733.

Zahm, S. H. *et al.* 1990. A case-control study of Non-Hodgkin's Lymphoma and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) in eastern Nebraska. *Epidemiology* **1**, 349–356.

ENDNOTES

1. Agricultural Health Study Information Packet for Advisory Panel Meeting, January 18–19, 1996 — Meeting Overview, NCI Summary, Biomarkers Studies,

Federal Government's Agricultural Health Study

NIEHS Summary, Validation Studies, USEPA Summary, IA Field Station Summary, NC Field Station Summary; Agricultural Health Study Information Packet for Advisory Panel Meeting, February 27–28, 1997 — AHS Study Update, NCI Summary, IA Field Station Summary, NC Field Station Summary, NIEHS Summary, USEPA Summary, Collaborative Efforts, Future AHS Activities.

2. All response rates are from Agricultural Health Study Information Packet for Advisory Panel Meeting, March 4–5, 1998 — NCI Progress Report.