

EXAMINATION OF NONRESPONSE BIAS IN THE UNIVERSITY OF MICHIGAN DIOXIN EXPOSURE STUDY

Olson, K¹, Sinibaldi, J¹, Lepkowski, J¹, Lohr-Ward, B¹, Ladronka, K¹

¹Institute for Social Research, 426 Thompson Street, Ann Arbor, Michigan 48104

Introduction

Scientific surveys are a commonly used tool to understand what is happening and why it is happening in a population of interest. An important feature of scientific surveys is the use of a probability sample of the population. In a survey with a probability sample, units are selected with known, nonzero probabilities and systematically approached for participation in the survey. The decision to participate or not participate is recorded for each unit of the population. Inferences based on probability sampling methods will be valid when all selected units respond to the survey request. When any proportion of the sampled units do not respond, inference from the survey may be at risk. The difference between estimates based on the respondents alone and estimates based on the full sample is called "nonresponse bias."

Estimating the effect of nonresponse bias on survey estimates requires mounting special studies to obtain information on those individuals who did not initially respond, known as the nonrespondents. One way to identify information on the nonrespondents is to ask them once again to participate in the study but use a different protocol-- a technique known as a nonresponse follow-up study. Although a nonresponse follow-up study is most informative about nonresponse bias when 100 percent of the nonrespondents reply, obtaining information about any of the nonrespondents can be useful for understanding whether the respondents and nonrespondents differ on the characteristics of interest.

The University of Michigan Dioxin Exposure Study used an area probability sample to examine the relationship between dioxins in the environment and dioxins in people's bodies. While response rates were high (74.3 percent, AAPOR RR3)¹, some selected households did not have eligibility determined or complete the questionnaire. This paper looks at whether the prevalence estimates from the questionnaire were affected by nonresponse bias.

Materials and Methods

Data used in these analyses are from the questionnaire designed for the University of Michigan Dioxin Exposure Study and the shortened questionnaire from the nonresponse follow-up. The questionnaire for the main interview was approximately one hour in length, and administered in-person by trained interviewers from the Survey Research Center at the University of Michigan. The nonresponse follow-up questionnaire was a shortened version of the main interview questionnaire, approximately 15 minutes in length, and administered by trained interviewers primarily over the telephone. Both questionnaires included extensive questions about possible exposure pathways, including residential history, occupational exposure, exposure through recreational activities, and consumption of meat, fish, and game. Basic demographic and health questions were also asked.

The protocol for recruitment for the nonresponse follow-up differed from the main study recruitment protocol² in the following ways. First, respondents to the nonresponse follow-up were not asked to provide serum, dust or soil samples. Second, the questionnaire was shortened dramatically, from approximately 60 minutes to approximately 15 minutes. Nonrespondents were mailed a \$20 bill in a (tailored, if possible) advance letter for the nonresponse follow-up, and were promised an additional \$40 after completion of the 15 minutes interview. Interviewers were permitted to use the telephone for identifying and selecting the respondent and for conducting the interview. The nonresponse follow-up for nonrespondents to the Fall 2004 data collection was conducted in January and February 2005; nonrespondents to the Spring 2005 data collection were contacted in October and November 2005.

Dioxin exposure study in Midland, MI

The UMDES obtained a 74% (RR3) response rate to the main questionnaire and a 50% (RR3) response rate among the nonrespondents to the shortened questionnaire, resulting in a 74% response rate overall. Thus, at least some information was obtained on 1504 individuals from the study populations, XX in the floodplain of the Tittabawassee River (590 from the main data collection /XX from the nonresponse follow-up), XX in surrounding areas of Midland and Saginaw Counties (XX/XX), and XXX in the control sites of Jackson and Calhoun Counties (359/XX).

The analyses for this paper are descriptive. Means and standard errors are calculated for each subgroup. Differences between the means for the subgroups are tested using t-tests and are considered significant at the $p < 0.05$ level. All analyses are weighted for differential probabilities of selection and adjustments due to nonresponse. Additionally, all analyses use imputed data to account for any missing data, "don't know" answers or refusals to answer questions. Standard errors are calculated to account for the complex survey design. All analyses were conducted using SAS 9.1.3.

The analyses compare the respondents to the main study ("respondents") to the respondents to the nonresponse follow-up study ("nonrespondents") across the five geographic regions of interest: the floodplain, near the floodplain, the aerial plume surrounding the Dow Chemical plant, other areas in Midland, Saginaw and Bay Counties, and the control site of Jackson and Calhoun Counties.

Results and Discussion

Analyses will include examining differences demographics such as age, sex, race, income, education, ownership of the property, length of residence on the current property, and household size. Health variables will also be of concern and these will include body mass index, smoking history, eligibility to give a blood sample, and breastfeeding.

Analyses of questions specifically relating to exposure pathways will include differences in dermal contact exposure through gardening, occupation, flooding of the property, serving in the Vietnam War, or recreational activities on or near the Tittabawassee River. Potential for exposure through ingestion is also of importance and will involve the analyses of the respondents' diets. In the condensed follow-up questionnaire, respondents were asked about their wild fish and game consumption in detail. This will allow comparison to the longer questionnaire used in the main data collection.

The analyses will identify differences between the respondents to the main study and those in the nonresponse follow-up. Since blood serum was not collected from the respondents in the follow-up, definite conclusions about the differences in dioxin levels in the two subgroups cannot be made. However, if well specified causal models are developed to predict blood dioxin levels using the variables found in the follow-up questionnaire, assumptions can be made about the existence and direction of the bias of the findings.

References

- 1 The American Association for Public Opinion Research. 2006. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 4th edition*. Lenexa, Kansas: AAPOR.
- 2 Lepkowski J, Olson K, Ward B, Ladronka K, Sinibaldi J, Franzblau A, Adriaens P, Gillespie BW, Chang SC, Chen Q, Demond A, Gwinn D, Hedgeman E, Knutson K, Lee SY, Sima C, Swan S, Towey T, Zwica L, Garabrant D *Organohalogen Comp* 2006 (forthcoming).