

AN EXAMPLE OF BIAS IN THE NAHMS SWINE '95 STUDY DUE TO QUESTIONNAIRE STRUCTURE

Bush EJ

USDA:APHIS:Centers for Epidemiology and Animal Health
555 S. Howes St, Fort Collins, CO

The USDA's National Animal Health Monitoring System (NAHMS) conducts national studies to provide scientifically-sound, statistically-valid information for education, research, policy development, and the overall improvement of animal industries. The probability sampling design used for the Swine '95 study permits the estimation of measured variables for the national population.

The first questionnaire, the General Swine Farm Report (GSFR), was completed on 1477 pork producers in 16 states. Interviews were conducted by enumerators for USDA's National Agricultural Statistics Service (NASS) between June 1 through June 23, 1995. Two additional questionnaires were completed on 418 respondents by State and Federal Veterinary Medical Officers between July 17, 1995 and January 17, 1996. All questionnaires were completed by face-to-face personal interviews.

National estimates based on the NAHMS Swine '95 study have been published in three parts [USDA, 1995; USDA, 1997]. The quality of the estimates are typically assessed from appropriate calculation of sampling variance, i.e. standard errors. However, non-sampling errors at data collection can also adversely effect the quality of estimates [Bush, 1995]. The interview consists of a dynamic interaction between the interviewer and respondent via a questionnaire; each of these three components may contribute to observational errors in an epidemiological study. This paper documents the presence of non-sampling error due to the questionnaire in the NAHMS Swine '95 national study.

In Section 4 of the GSFR, question 12 asked
“During the last 12 months, how many times did this hog and pig operation test:
a. Ground water (for nitrates or pathogens)?
b. Nutrient content of manure (such as nitrogen level)?
c. Air quality (such as ammonia or hydrogen sulfide levels)?”

National estimates based on this question from the GSFR are presented (without standard errors) in Table 1. They were not published in the Swine '95 Part I report due to suspicions that they were biased. Environmental testing of water, manure, and air are infrequently done by producers and a distribution heavily skewed towards zero was expected. The elevated estimates for environmental testing three times a year suggested coding errors due to questionnaire structure.

Test	Number of times tested				
	0	1	2	3	>3
Ground water	75.5	11.3	1.0	11.9	0.2
Nutrient Content	80.2	3.9	1.0	14.8	0.2
Air quality	82.4	1.3	0.4	15.3	0.6

Table 1. During the 12 months prior to the June 1995 interview, percent of operations by number of times the following were tested.

Due to suspected bias in these estimates, they were withheld from publication in the Swine '95 Part I report and the question was reasked in the Swine Health Report (SHR) using the exact same wording. The interview for the SHR was conducted approximately five months after the GSFR interview. The national estimates based on respondents to the SHR are presented in Table 2.

Test	Number of times tested				
	0	1	2	3	>3
Ground water	84.6	11.9	2.3	0.6	0.6
Nutrient Content	92.2	5.8	1.5	0.2	0.3
Air quality	97.8	1.6	0.4	0.2	0.0

Table 2. Percent of operations performing tests of groundwater, nutrient content of manure, and air quality during the previous 12 months by number of times tested.

Standard errors for both sets of estimates were calculated based on methods for complex surveys [Dargatz]. Estimates for testing 0 and 3 times in last 12 months were statistically different between the two questionnaires, based on non-overlap of 95% confidence intervals, while no significant difference existed for other levels. The estimates are not expected to be identical since they reflect different time periods, different interviewers, and different populations (those with at least one hog for the GSFR versus those with at least 300 finishers for the SHR).

The significant difference in those indicating that no environmental testing had been done may reflect the influence of secular trends, interviewer effect, and random variation. However, the most likely explanation is that those previously coded as 3 on the GSFR were now coded as 0 on the SHR.

The significant drop in the percent of producers indicating they conducted environmental testing three times in the previous 12 months supports our suspicion that coding errors occurred in the GSFR. Question 12 on the GSFR appeared as the fourth question out of five on page 11. With the exception of question 12, all questions on this page were yes/no questions which are coded as 1 for Yes and 3 for No. This is standard coding mechanism for NASS enumerators.

To confirm that GSFR respondents coded as a 3 were really misclassified 0's (no testing), data from the SHR on the same respondents were evaluated. The specific hypothesis tested was that GSFR respondents coded as 3 were more likely to be recoded as a 0 on the SHR than the GSFR respondents which were coded as a 1, 2, 4 or more. Those coded as a 0 on the GSFR were excluded from this analysis.

GSFR	Recoded on SHR as 0	
	Yes	No
1,2, 4+	55	48
3	32	10

Table 3. Ground water.

GSFR	Recoded on SHR as 0	
	Yes	No
1,2, 4+	42	43
3	40	6

Table 4. Nutrient content of manure.

GSFR	Recoded on SHR as 0	
	Yes	No
1,2, 4+	22	15
3	53	3

Table 5. Air quality.

A significant difference based on chi-square testing ($p < 0.01$) on all three tables confirms that estimates on frequency of environmental testing based on GSFR were biased due to misclassification of respondents doing no testing. This bias was due to the location of an open-ended question with a discrete number of responses in the midst of a series of yes/no questions. This finding underscores the importance of non-sampling errors in epidemiological surveys.

References

1. USDA:APHIS:VS-CEAH. Swine '95 Part I: Reference of 1995 Swine Management Practices. 1995. Number: #N186.995
2. USDA:APHIS:VS-CEAH. Swine '95 Part II: Reference of 1995 U.S. Grower/Finisher Health & Management Practices. 1996. Number: #N201.696
3. Bush EJ, Cowen P, Morrow WEM, Dickey DA, Zering KD: Evaluation of non-sampling errors in the U.S. National Swine Survey. *Prev Vet Med* 1995; 22: 155-168.
4. Dargatz DA: Analysis of survey data. *Prev Vet Med* 1996; 28: p. 225-237.