AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 97, No. 4

CHARACTERISTICS OF RESPONDENTS AND NONRESPONDENTS TO A QUESTIONNAIRE FOR ESTIMATING COMMUNITY MOOD[1]

GEORGE W. COMSTOCK AND KNUD J. HELSING

(Received for publication December 8, 1972)

Comstock, G. W. (Training Center for Public Health Research, Johns Hopkins School of Hygiene and Public Health, Box 2067, Hagerstown, Md. 21740) and K. J. Helsing. Characteristics of respondents and nonrespondents to a questionnaire for estimating community mood. *Am J Epidemiol* 97: 233–239, 1973.—In a study of community mood in Washington County, Maryland, a representative sample of adult residents was selected for interviews. Among this sample, a total of 649 persons had participated in a previous nonofficial census and for this group, similar information was available both for the 571 who were interviewed and the 78 who refused. Characteristics associated with refusals were poor education, residential stability and small households. There was also suggestive evidence that young adults and males were more likely to refuse. Urban-rural residence, cigarette smoking and frequency of church attendance did not appear to be related to refusal.

health surveys; respondents and nonrespondents, characteristics; sampling studies

Much of our knowledge about national and community health has been gathered by means of door-to-door interviews. This information, like most observational data on human populations, can be affected considerably by response bias. It is surprising, therefore, to find that little has been learned about differences between persons who respond to interviews and those who do not. This lack of knowledge apparently results from the understandable difficulties in obtaining similar information concerning respondents and nonrespondents in most situations. In contrast, much more is known about factors affecting response to a variety of procedures involving physical or laboratory examinations, either as a result of preliminary or subsequent interviews (1–4). Among persons who are interviewed, it is simple to compare those who accept a subsequent examination with those who do not. It is not so easy to study the characteristics of persons who refuse to be interviewed.

One method of estimating whether persons who respond to an interview differ from the population they are intended to represent is to compare certain characteristics of the respondents with official census information for the same geographic entity and for approximately the same time. This approach was used in the first of a series of morbidity surveys that culminated in the present National Health Surveys (5). In a

[1] From the Community Mental Health Epidemiology Program, Center for Epidemiologic Studies, National Institute of Mental Health; and the Training Center for Public Health Research and the Department of Epidemiology, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Maryland 21205.

Reprint requests to Training Center for Public Health Research, Box 2067, Hagerstown, Maryland 21740.

This study was supported in part by Contract HSM 42-71-32 from the National Institute of Mental Health, and by Research Career Award HL 21,670 from the National Heart and Lung Institute.

study of illnesses in a sample of households in Hagerstown, Maryland from 1922 to 1924, the age distribution of the study population was compared to that of the population of Hagerstown obtained in the 1920 census. More than two decades later, a nationwide study of medical costs and health insurance and a morbidity survey in Hunterdon County, New Jersey used the same approach, but looked at several socioeconomic indices in addition to age (6, 7).

All three reports commented on the close agreement between the study and Census populations. However, the similarity of such distributions can be more apparent than real. Direct visual comparisons of percentage distributions lack sensitivity. If the ratios of the percentages of the study and Census populations in each age group are compared for these studies, it is quickly apparent that each of the three study populations had a relative excess of school age children and a relative deficit of young adults. For example, in the Hagerstown study, persons five to 14 years of age comprised 23.6 per cent of the study population, but only 18.9 per cent of the Census population, a ratio of 1.25 to 1. Young adults 20 to 34 years of age made up 22.3 per cent of the study population compared with 27.2 per cent of the Census population, a ratio of 0.82 to 1. In the Hunterdon County study, persons with a high school education or better were over-represented compared to the Census population, while in both of the later studies persons with less than a grade school education were under-represented, as were farm women and farm laborers. In Hunterdon County, persons classified as professionals and proprietors were better represented than other occupational groups, but this was not true in the nationwide survey of medical costs. An unavoidable drawback to this method of characterizing nonresponse is that differences from the Census population are usually taken to imply inadequate response whereas they may actually result from sampling and nonsampling errors or from changes in the population since the Census was taken.

An entirely different approach was used in a survey of a sample of households in the Washington Heights district of New York City, in which interviews were sought from occupants of 4500 dwelling units during 1960–1961 (8). By the first cut-off date, 77 per cent of the interviews had been obtained. Of those remaining to be interviewed, approximately 550 had refused and 450 had not yet been found at home. A third of these "hard-to-obtain" groups was assigned to a special team, who obtained interviews from 60 per cent of them. Greater success was had with persons who had not been found at home initially than with those who had refused. An indication of the potential effects of nonresponse was then derived by comparing the characteristics of persons interviewed initially with those who were previously unlocated and with those who had refused initially. Of 200 items available for such comparisons, only 15 showed differences. There were no differences by sex, race, religion, occupation or family income. The group who had been hard to locate had fewer children under 16 and fewer adults under 35, and had more one-person families and more families with an adult over 65 than the initial interviewees. The initial refusals tended to consist of families in which the adults lacked a high school education, worked full-time or were all over 65. Both groups of initial nonrespondents showed residential stability, having a high proportion who had lived at the same address for more than 15 years. Although the method used in this study is free of some of the problems involved in matching the characteristics of study and Census populations, it yields no information about persons who refuse to respond even after intensive efforts.

The Human Population Laboratory of the California Department of Public Health also examined potential nonresponse bias after successive stages of follow-up ef-

forts among a population with known demographic and socioeconomic characteristics (9). The baseline information had been obtained from 8083 adults in Alameda County by means of household interviews. At the conclusion of each interview, questionnaires were left for all adults with the request that they be returned by mail. Only 42 per cent responded. After a mailed reminder, the returns rose to 59 per cent. The rate of return was further increased to 70 per cent by a telegraphic request and to 86 per cent by all of the preceding efforts plus a personal visit. Comparisons of the characteristics of the persons responding before the personal contact stage with nonrespondents showed a significant excess of the following among the nonrespondents: persons over 65 (the only reported age categories were under and over 65), Negroes, service workers and laborers, and persons living in crowded homes or in homes with no telephones. Personal visits increased the responses among persons with the preceding characteristics but at the cost of causing over-representation of categories of persons who were easily found at home compared to their true representation in the base population. There was no report of multifactorial analyses aimed at discovering which demographic and socioeconomic characteristics were most closely linked with nonresponse.

A similar opportunity to study some characteristics of nonrespondents to an interview occurred in Washington County, Maryland, nearly 50 years after the pioneering morbidity survey conducted by Sydenstricker in this area (5). A questionnaire designed to measure mood with special emphasis on depressed feelings was administered every week to a different random sample of adult residents. Matching this sample to persons identified in a nonofficial census of the entire county done in 1963 yielded a population in which current respondents and nonrespondents could be compared on the basis of previously ascertained characteristics.

Materials and methods

In the summer of 1963, a nonofficial census of Washington County, Maryland was carried out by the Johns Hopkins Training Center for Public Health Research, Washington County Health Department and the National Cancer Institute (10). Information obtained at that time included race, sex, age, years of schooling, smoking history and frequency of church attendance. A recheck of approximately 10 per cent of the county indicated that better than 98 per cent of the households had been enumerated.

In 1971, the list of dwelling units obtained in the 1963 census was brought up to date by adding units constructed since that time. The new units were identified from building permits, tax records and Health Department records. This list provides the sampling frame for a number of studies, including an assessment of community mood through interviews obtained each week from a random sample of residents over the age of 18 years. With a separate random start for each week's sample, 32 or 33 dwelling units were systemically selected. Because some units were currently vacant or had been demolished, this number yielded a sample of approximately 30 occupied units each week. The selected households were assigned to four pleasant adult female interviewers, each of whom was responsible for a quadrant of Washington County and the City of Hagerstown. Interviewers were assigned to a new quadrant every 11 weeks. After locating someone at home in the assigned household, the interviewer first ascertained the composition of the household, listing all persons over 18 years of age in a specified order. If more than one person was listed, further sampling was necessary, and the person to be interviewed was chosen according to random numbers in sealed envelopes. If the selected person was at home, an interview was requested at that time; if not, an appointment

TABLE 1
Derivation of the study population

Initially selected households		1206
Vacant or torn down	106	
Households available for interview		1100
Stage 1 exclusions:		
No one home during 4-week period	14	
Insufficient information to select a respondent	69	
Households with potential respondent identified		1017
Stage 2 exclusions, after respondent selected:		
Respondent could not be contacted in 4-week period	3	
Incomplete interview	2	
Difficulty with language or comprehension	6	
Respondent not in 1963 census	301	
Respondent under 25 years of age	56	
Study population		649
Interviews obtained		571
Refusals		78

was made for a later visit. In all instances, the purpose of the study and the general nature of the questions were explained to the subjects, who were scrupulously told that they were free to answer as many or as few questions as they chose.

Nonresponses could occur at either of two stages in this interviewing procedure. Of the 1206 dwelling units selected during the first 37 weeks of this study, 1100 appeared to be occupied and still in residential use. At the first stage, it was not possible to find anyone at home in 14 dwelling units before the four-week deadline for obtaining interviews could be met, and in 69, the first person contacted refused to identify the household members, so no selection of potential respondents could be made.

At the second stage, three adults selected for interviews could not be contacted within the specified time, two interviews were started but not completed, and six were not interviewed for other reasons such as language difficulty, senility or illness of the respondent. This left a total of 1006 identifiable potential respondents.

Records of these potential respondents were matched to listings from the 1963 census. A total of 705, approximately 70 per cent, of these individuals were identified as having participated in the census, and information from the census was transferred to their current study records. Fifty-six persons under 25 years of age at the time of interviewing were excluded from the analysis because personal histories were not obtained from persons under 17 years of age in 1963. This left a total of 649 persons who were identified in the census, who were 25 years of age or older at the time of interview, and who either were interviewed or refused to be interviewed. Table 1 shows the exclusions from the initial sampling selections with the reasons for exclusions from this study.

RESULTS

Of the 649 persons in the study population, 78 or 12.0 per cent refused to be interviewed. The refusal rate was essentially the same for all interviewers. Although there were fluctuations in the frequency of refusals from week to week, no long-term time trends were noted over the period of this analysis which extended from December, 1971 to August, 1972.

Refusal rates according to various personal and demographic characteristics are shown in table 2. In addition to the crude refusal rates, which show the situation as it was actually encountered, the results are also shown as rates adjusted for the effects of the other variables by the binary multiple regression method of Feldstein (11). The adjusted rates indicate the relative importance of each factor when the confounding effects of other (possibly correlated) factors are removed. The adjusted rates are also shown in figure 1.

Males were more likely to refuse than females, a difference that did not quite

reach a statistically significant level after adjustment.

The crude refusal rate increased with increasing age. However, when the effects of other variables were taken into account, a dramatic reversal of the trend was noted, the adjusted refusal rate being 16.4 per cent among persons 25 to 39 years of age compared to 10.4 per cent among those over 55, and 10.9 per cent among those between 40 and 54 years of age.

There was a decreasing rate of refusals with the number of persons currently in the household, both in the crude rates and adjusted rates. The Cochran test for linear regression shows the trend to be significant at $p < .025$.

The likelihood of encountering a refusal was not significantly different in city, suburbs and countryside. Residential stability, however, was significantly related to the probability of a refusal. Persons who were currently at their 1963 addresses were much more likely to refuse than persons who had remained in the county but had changed addresses.

The most striking association was with the amount of education of the selected person. Both the crude and adjusted rates showed that persons with at least a high school education were least likely to refuse and that persons with less than nine years of schooling were most likely to refuse.

Almost no association of refusals was found with cigarette smoking habits in 1963. Infrequent church attenders were slightly but not significantly more likely to refuse than persons who went to church at least once a week.

TABLE 2

Percentage of persons who refused to answer questionnaire, by personal and demographic characteristics

Characteristic	No. of persons contacted	Refusals		
		No.	%	Adjusted %*
Total	649	78	12.0	
Sex:				
Male	282	41	14.5	14.9
Female	367	37	10.1	9.8
Age in 1972:				
25–39	159	14	8.8	16.4
40–54	213	23	10.8	10.9
55+	277	41	14.8	10.4
No. of people in household in 1972:				
1	130	20	15.4	16.0†
2	230	33	14.3	13.8
3+	289	25	8.7	8.9
Residence in 1972:				
Urban	239	25	10.5	10.6
Surburban	76	11	14.5	16.5
Small town	70	11	15.7	15.2
Rural	264	31	11.7	11.2
Living in same house in 1963 & 1972:				
Yes	394	57	14.5	14.6‡
No	255	21	8.2	8.1
Years of schooling in 1963:				
0–8	221	43	19.5	19.3§
9–11	161	17	10.6	10.9
12+	267	18	6.7	6.7
Cigarette smoking in 1963:				
Yes	253	30	11.9	11.7
No	396	48	12.1	12.2
Church attendance in 1963:				
Once a week or oftener	260	25	9.6	10.1
Less than once a week	389	53	13.6	13.2

* Adjusted for all other variables in table by Feldstein's method of multiple binary regression.
† $p < .025$ when tested for trend.
‡ $p < .025$.
§ $p < .001$.

DISCUSSION

This investigation shares with the Washington Heights survey the defect that not all refusals could be studied. There was no way to tell if hardcore refusals in Washington Heights differed in other respects from refusals who later agreed to be interviewed. Nor was it possible in Washington County

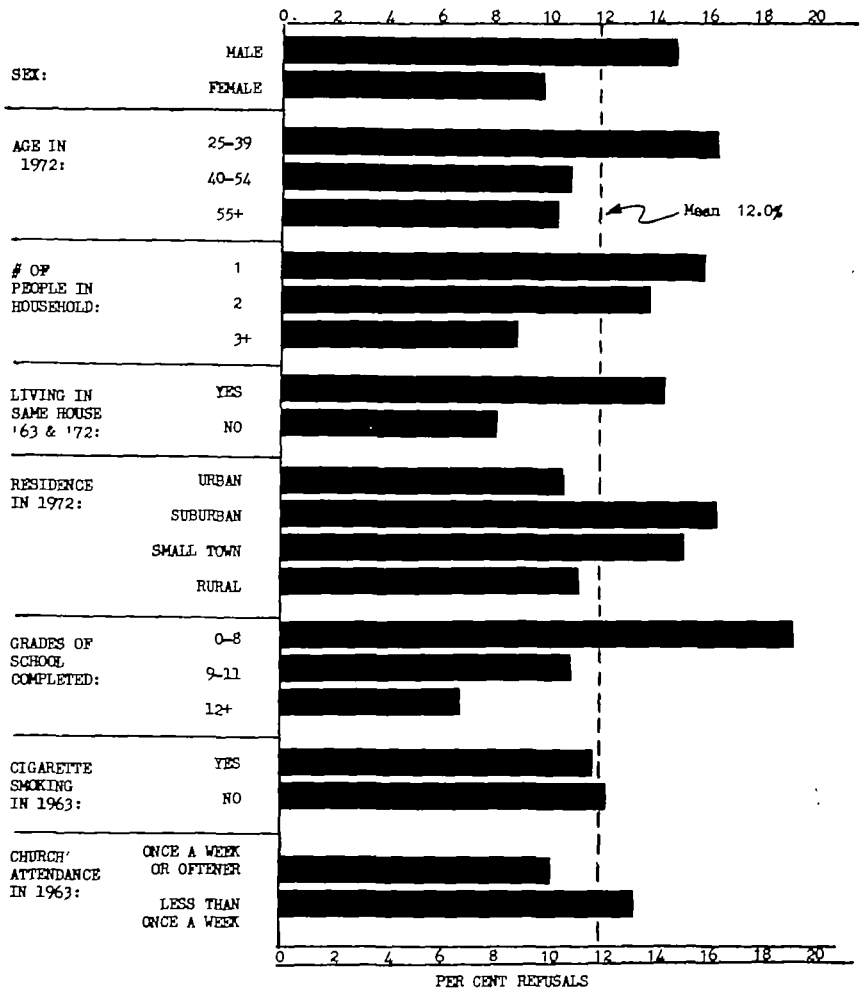

FIGURE 1. Adjusted refusal rates by personal and demographic characteristics

to ascertain if persons living in households where it was not possible to obtain a listing of all adults differed from persons who were contacted but did not wish to be interviewed.

On the other hand, the present study has certain methodologic advantages over its predecessors. Comparisons with a census population are much more meaningful if individual study records can be matched to the corresponding census records. For official censuses, such matching can only be done by the Census Bureau, because of their entirely appropriate insistence on confidentiality. Because we know so little about characteristics of persons who refuse to respond to interviews and because interviews are so commonly used to provide vitally necessary knowledge, it would be highly desirable to take advantage of the wealth of socioeconomic information gathered in Decennial Censuses. This could probably be done if persons undertaking interview studies at the time of an official census were to make arrangements for the Census Bureau to do the matching and to produce the tables for what could be a thorough study of the socioeconomic and demographic characteristics of nonrespondents.

Three characteristics of refusals stood out in the present analysis. Education seemed to be the most important factor, with resi-

dential stability and number of people in household also being significantly important. Confidence in the potential importance of these attributes is increased by the fact that all of them also appeared to be important in other studies.

In studies of response to interviews, it is desirable to show the crude and adjusted rates. Crude rates reveal the situation as it exists in the particular sample and hence are of considerable value both for local administration and for descriptive purposes. Adjusting the rates for other pertinent variables comes closer to indicating the factors underlying response and refusal. For example, the crude rates in this study indicated that increased effort should be devoted to gaining the cooperation of older people. However, examination of the adjusted rates suggested that the most important determinant of response in this study was education. The high crude refusal rate of older persons appeared merely to reflect the fact that a high school education was not nearly so common 30 years ago as it is today.

It is not easy to know how to apply knowledge about response to interviews. In the first place, we have little such knowledge to apply. And in the second place, we know even less about which tactics will prove most successful under various conditions. Possibly interviewers should be selected for characteristics that correspond to those of the potential subjects. This cannot be done on an individual basis, of course, because the subject's characteristics cannot be known prior to interview in most situations. However, the selection of interviewers could be based on the overall characteristics of an area if this procedure were shown to be beneficial. Studies of response rates to interviews and of methods to improve them are long overdue.

References

1. Cobb S, King S, Chen E: Differences between respondents and nonrespondents in a morbidity survey involving clinical examination. J Chron Dis 6: 95–108, 1957
2. National Center for Health Statistics: Factors related to response in a health examination survey, United States, 1960–1962. Vital and Health Statistics, Public Health Service Publication No. 1000, Series 2, No. 36. GPO, Washington, 1969
3. Naguib SM, Geiser PB, Comstock GW: Response to a program of screening for cervical cancer. Public Health Rep 83: 990–998, 1968
4. Cochrane AL, Cox JG, Jarman TF: Pulmonary tuberculosis in the Rhondda Fach. An interim report of a survey of a mining community. Br Med J 2: 843–853, 1952
5. Sydenstricker E: A study of illness in a general population group. Hagerstown morbidity studies no. I: The method of study and general results. Public Health Rep 41: 2069–2088, 1926
6. Commission on Chronic Illness: Chronic illness in the United States. Vol. III. Chronic illness in a rural area, the Hunterdon Study. Cambridge, Harvard University Press, 1959, pp 62–73
7. Anderson OW, Feldman JJ: Family medical costs and voluntary health insurance: A nationwide survey. New York, McGraw-Hill Book Co, 1956, pp 223–232
8. Loewenstein R, Colombotos J, Elinson J: Interviews hardest-to-obtain in an urban health survey. Milbank Mem Fund Q 47 (No 1, Part 2): 195–200, 1969
9. Hochstim JR, Athanasopoulos DA: Personal follow-up in a mail survey: its contribution and its cost. Public Opinion Quarterly 34: 69–81, 1970
10. Comstock GW, Abbey H, Lundin FE Jr: The nonofficial census as a basic tool for epidemiologic observations in Washington County, Maryland. *In* The Community as an Epidemiologic Laboratory. Edited by Kessler II, Levin ML. Baltimore, Johns Hopkins Press, 1970, pp 73–97
11. Feldstein MS: A binary variable multiple regression method of analyzing factors affecting perinatal mortality and other outcomes of pregnancy. J Roy Stat Soc 129 (Series A, part 1): 61–73, 1966