A SAMPLE OF DEVELOPMENTS FROM SAMPLING PROJECTS OF THE U. S. NATIONAL HEALTH SURVEY

By: Walt R. Simmons, National Health Survey

It was just a little over two years ago that the National Health Survey Act became a part of our laws. The act authorized the Public Health Service to develop and carry out a continuing program of statistics on the illnesses, injuries, impairments, and related characteristics of the population. Need for this legislation had been long felt, and carefully studied. The history of experimentation and study in this area, and delineation of many of the purposes and objectives of the new law have been reported elsewhere.[1] They will not be reviewed here except to note that the health of a Nation's people is possibly its greatest asset, and that measurement of levels and trends in health matters is essential to the efficient promoting of better health.

Within the Public Health Service the Surgeon General established the U. S. National Health Survey as the organization immediately responsible for conducting the program. Enabling legislation specifically authorized a three-phase program to accomplish its general intent. The first of three phases is a continuing survey of the general civilian population to provide statistics on the incidence and prevalence of acute and chronic illnesses; on numbers of accidental injuries and impairments; on volume and kind of disabilities; on use of medical, dental, and hospital facilities; and on a variety of allied topics.

The second phase consists of special studies, designed to complement the continuing survey with selected determinations which are narrower in scope than the more general undertaking. The special studies may relate to a particular aspect of specified health statistics, to a type of measurement which cannot be made in the more general survey, or to other features of the health status of people.

The third phase of the program is of particular note to members of a statistical association. It is of course common that any substantial ongoing activity will devote a part of its energy to improvement in operation. In the case of the Health Survey, however, the Congress directed that methods and survey techniques for securing statistical information on health be studied, with a view toward their continuing improvement.[2]

As had been intended, approximately the first 12 months following the passage of the Health Survey Act were devoted to establishing the nucleus of an organization, securing staff, initial planning, testing,[3] and as urged by the Act, to making contracts with other agencies whose facilities and capacities would be especially useful in fulfilling objectives of the program.

Now the National Health Survey has completed recently its first full year as an operating organization. It is the purpose of this paper to offer a sampling of that first year's activity. This is sampling somewhat in the nature of a cook sampling the soup with a spoon; or the commuter's survey of current American and world affairs by scanning his favorite morning paper. The report is not intended as a complete summary of the year, nor yet a scientific random sample of such a summary. It is hoped that identification of a few developments in each of the three phases of the program, with brief comments on those developments, will give some impression of the course the National Health Survey has been taking.

Findings From the Continuing Survey of the Population

It was decided early that a principal part of the continuing survey should be an area-type probability sample of households. In accordance with requirements and specifications set by the Public Health Service, arrangements were made for the Census Bureau to design the sample and to plan and conduct the household interviewing.

The household survey follows a highly stratified multistage probability design which permits a continuous sampling of the civilian population of the United States. The first stage of the initial design consists of an area sample of 372 from among 1,900 geographically defined Primary Sampling Units (PSU's) into which the continental United States has been divided. A PSU is a county, a group of contiguous counties, or a Standard Metropolitan Area. Subsequent stages of selection result in the choice of ultimate units of compact segments containing an expected six households each.

The central design relates to a sample of the population to be interviewed over the span of one year. In the year just passed, approximately 115,000 persons in 36,000 households over the Nation came into the sample. A striking and distinguishing feature of the survey is its continuous nature. About 700 households are interviewed each week. This block of households is itself a representative sample of the civilian population of the United States. Thus the design permits a continuous measure of characteristics of high incidence or prevalence, and through consolidation of, say, 13 or 52 weekly samples, more detailed analysis and treatment of less common characteristics and smaller categories.

The interview is rigidly structured, consisting in its first year of 40 items for identification of persons and households and their socio-economic characteristics, and of 54 questions concerning health conditions of respondents.

Both manual and electronic editing and tabulating are performed at the Census Bureau under Public Health Service specifications. The formal estimation process is a two-stage ratio method, sample results being controlled by residence, age, sex, and color of the population as determined by the 1950 decennial census and more recent independent estimates.

Five principal reports[4-8] of substantive findings from the household interviews have been released. They cover physician visits, dental visits, persons injured, disability, and a set of subjects termed "Selected Survey Topics." The first four of these are provisional, being based on either 13 or 26 weeks of interviewing. The fifth is in the nature of a partial summary of data collected during the first full year's interviews. A sixth report on acute conditions has been completed and will be released in a few days. Reports on hospitalization and on impairments are scheduled for publication before spring.

Copies of these reports are available. Here we shall note only a very few findings, chosen principally to illustrate types of information being released.

In a preliminary report on volume of physician visits it was observed that during the third calendar quarter of 1957 residents of the United States visited the doctor at an average rate equivalent to five times a year. Two thirds of all visits involved diagnosis or treatment, with one third representing preventive care and miscellaneous purposes. Over-all global figures of this type were issued initially to help sketch the general dimensions of various health problems, even though it was recognized that they rested on rather small samples, might have seasonal biases—since they were based on experience in but one calendar quarter—and would have other shortcomings. In some measure all findings from the survey will be considered tentative and provisional, pending subsequent evidence, although clearly reports based on longer periods of sampling and those treating phenomena of high frequency and relative stability will be more nearly definitive than other types of reports. Preliminary reports will be superseded with more reliable data as time and resources permit. For example, that first estimate of number of physician visits per person per year already has been replaced with an estimate based on 52 weeks of interviewing. The estimated rate was revised from a little under five visits per person per year to a little over five visits.

One figure picked up by the press was another provisional estimate from one quarter's interviewing. It indicated that there are 22 million edentulous persons in the United States. Reporters phrased this finding, "By Gum: 22 Million with Nary a Single Tooth."

A beginning has been made in assembling information on injuries and accidents—their types and distribution. Data from the first year's interviewing show nearly 50 million persons being injured annually. Injury here refers to bodily hurt or damage sustained from external causes either accidental or purposeful. Survey reports show some 40 percent of all injuries occurring in the home, about 15 percent at work outside the home, another 10 percent involving motor vehicles, and the remainder otherwise classified.

First publications on injuries point up a feature which will be a part of many Survey statistics. This is the severity of the condition. It will be recognized that ailments come in many degrees, and that the statistics on illnesses or injuries are the resultant not only of chosen definitions but also of the entire measurement procedure. Without going into this matter at length, it is noted that Survey publications include what it is hoped are sufficiently detailed descriptions of statistics to permit the reader to acquire a reasonably clear understanding of just what it is that has been measured. In the case of the above numbers on injuries, for example, a condition which might be a candidate for recognition as an injury is not so counted in survey publications unless it involved medical attendance or at least one full day of restricted activity of the person hurt.

Incidentally only about 50 percent of the possible-injury conditions picked up by interviewers have met the criteria of medical attendance or one day of restricted activity. A further sidelight might be noted in this connection. Statistics on injuries grow out of a sequence of questions which locate the injury as occurring in the first week prior to interview or in the second week prior to interview. If there be no differential bias of recall for the two weeks, one would usually expect about 50 percent of reported injuries to occur in each of the two weeks. Sample checks show experience is consistent with this hypothesis. For lesser injuries not meeting the severity test, however, more than 75 percent occur in the first previous week, thus suggesting a rapid falling off in recall for these minor hurts as the period of recall lengthens.

To most of us it is the human and personal aspects of illnesses and impairments that are their dominating characteristics. But survey reports underline also the social and economic impact of ill health. Data from 52 weeks of interviewing produced an estimate of $3\,^{1}/_{3}$ billion man-days of restricted activity including $1\,^{1}/_{3}$ billion bed-days for the civilian noninstitutional population for the year ending last June. This is an average of 20 days of restricted activity and almost 8 bed-days per person per year. Another significant figure is 600 million days lost from work by persons 17 years old and over.

Technical Developments in the Continuing Survey

As implied earlier, it is a policy of the National Health Survey to prepare fairly extensive and detailed descriptions of its procedures and statistics. The first reason for this policy is one of self-discipline—it might even be called a selfish reason: it is simply that the exercise of reducing concepts and procedures to written form promotes a more precise understanding by the Survey staff of what is being attempted. But there are two other reasons for the policy. The first is that it permits consumers to know the data much better, and consequently to make better use of the data. The second is the belief that a widespread dissemination of information concerning Survey methods is very likely to lead to critical review, re-evaluation, and improvement of Survey products.

In line with this policy, each substantive publication issued thus far by the Survey has carried appendices covering such matters as definitions, sample structure, data collection, response rates, and sampling variances. More comprehensive treatment of these and other matters is being offered in a separate series of technical publications,[1,9,10] three of which have been issued thus far. The first of these papers dealt with the purposes, objectives, origin, and program of the Health Survey. The second is a detailed description of the statistical design of the household survey. The third is an exposition of the concepts and definitions employed in the household survey.

The initial household-survey design gave unusual emphasis to the securing of separate estimates for a large number of geographical sectors of the country. This course was taken in order to provide data which would detect the highlights of geographical differentials, even though it was apparent that the scale of operation was such that sampling errors for individual geographic sectors would be very substantial. Three factors have led to a design modification which is to be introduced the first week in January. These factors are (1) trends in consumer desires, and especially as expressed by advisory bodies, toward greater emphasis on national figures including somewhat finer detail in subject matter; (2) after 18 months of collection, the existence of data which can be tabulated by relatively small geographic sectors and thus perhaps permit a partial realization of the earlier demand for such information; and (3) the experience in operation of the survey which has pointed the way toward certain efficiencies. At a later date descriptions of the design modifications—which do not change the household-interview survey in any of its fundamental structural features—will be prepared. There are noted here just three leading features of the changes.

1. In line with the trend of consumer interest, resources have been reallocated to give greater emphasis to statistics of national totals based on 52 weeks of interviewing. But estimates still can be prepared for 12 geographic regions and for 4 different degrees of concentration of population.
2. The number of Primary Sampling Units (PSU's) has been changed from 372 to 500 in accord with the evidence of data on unit costs. Some PSU's will appear in the sample each week; some only once a year; the average number of appearances in a year for a single PSU is six. Each week's sample will contain some 50-60 PSU's and each quarter's sample approximately 400 PSU's. Each week's sample continues to be a representative sample of the population.
3. These two changes and others of lesser impact will yield a 25-percent reduction in variance for the new design over the old, with respect to a typical National Health Survey national statistic based on a full year's interviewing.

All that has been said thus far, on both methodological and substantive matters has given primary attention to results based on interviewing over a year or perhaps a calendar quarter. And indeed that emphasis is in accord with planning and policy of the Health Survey. In the very first weeks of operation of the household survey, an event occurred which, coupled with basic survey design, has led to experimental work in a somewhat different area.

The event was the epidemic of Asian influenza. The design feature was the fact that each week's interviewing in the National Health Survey is a representative sample of the population. Since there was urgent demand for information on the course of the epidemic, many people felt the new Survey should attempt to trace that course on a weekly basis.

It should be recalled that there are only about 700 households in each weekly sample. The Health Survey had never expected to attempt a weekly report on any topic. But Asian influenza was thought to be a phenomenon of very high incidence and desire for statistics on the topic was insistent. Accordingly, there was some hurried investigation of the possibilities. Since the investigation was sufficiently promising, the Public Health Service decided to issue weekly provisional releases on number of new cases of acute upper respiratory diseases (including pneumonia and influenza), and average number of persons in bed each day from such diseases. These reports were issued throughout the epidemic and then were extended to provide weekly data for the period from July 1957 to

May 1958.[11] At the peak of the epidemic there were nearly 12,000,000 new cases a week, and more than 6,000,000 persons were in bed.

The experience with what in shop talk came to be known as the Flu Data has led to exploration of other possible useful features of weekly data from the Health Survey. These explorations have not yet resulted in definite conclusions, although they may be characterized perhaps as encouraging. It is likely that any weekly estimates which may be produced would be useful on a retroactive, rather than a current basis.

The statistical features of this investigation and particularly the problems of estimating population parameters and sampling variances are intriguing. The present paper will not be diverted to those problems. One or two aspects may be noted briefly, however, not only for their intrinsic interest, but because they suggest features of the Household Survey which have implications outside the matter of estimates based on one week's interviewing.

For several types of conditions persons each week are asked in the Health Survey for onset of illness: (1) "Last week or the week before," and (2) separately for "Last week" and for "week before." Thus if α_i is the true measure of incidence for week i for a particular type of illness, two primary estimates of α_i are

α'_i, the estimate for "last week" obtained from interviewing in week (i+1), and

β'_i, the estimate for "week before" obtained from interviewing in week (i+2).

A number of other estimates may be derived from these primary estimates. Among those to which special attention has been given in our investigations are derived estimates u'_i and x'_i defined by the following equations.

(1) $u'_i = \frac{1}{2}(\beta'_{i-1} + \alpha'_i)$, and

(2) $x'_i = \frac{1}{2}(u'_i + u'_{i+1})$.

The latter quantity x'_i has several attractive features as an estimator of α'_i. It is a weighted average of four primary weekly estimates, two of which referred originally to week i, one to week (i-1) and one to week (i+1). The estimate x'_i comes from <u>two</u> weeks of interviewing in weeks (i+1) and (i+2), rather than from a single week.

These characteristics give x'_i a relative stability as compared with several other possible estimates. Specifically, the variance of x'_i is for all items less than ½ that of α'_i and for some items is a much smaller fraction of the variance of α'_i.

The statistic x'_i is a biased estimate. Its expected value being $\alpha_i + 1/4(\alpha_{i-1} + \alpha_{i+1} - 2\alpha_i)$. It is apparent, however, that the form of the bias is such that bias will be small for most statistics in most periods, except possibly for the trend turning point of a very rapidly changing phenomenon. Even in such instances, the bias is not a cumulating one, and is almost certain to be minor over a period of several consecutive weeks. Figure 1 reflects the interviewing and estimation scheme just described. Experimentation with weekly data continues. It is expected that conclusions on technical characteristics of these data will be reached shortly. Attention will turn then to more specific consideration of whether and how they can make an informational contribution to the program.

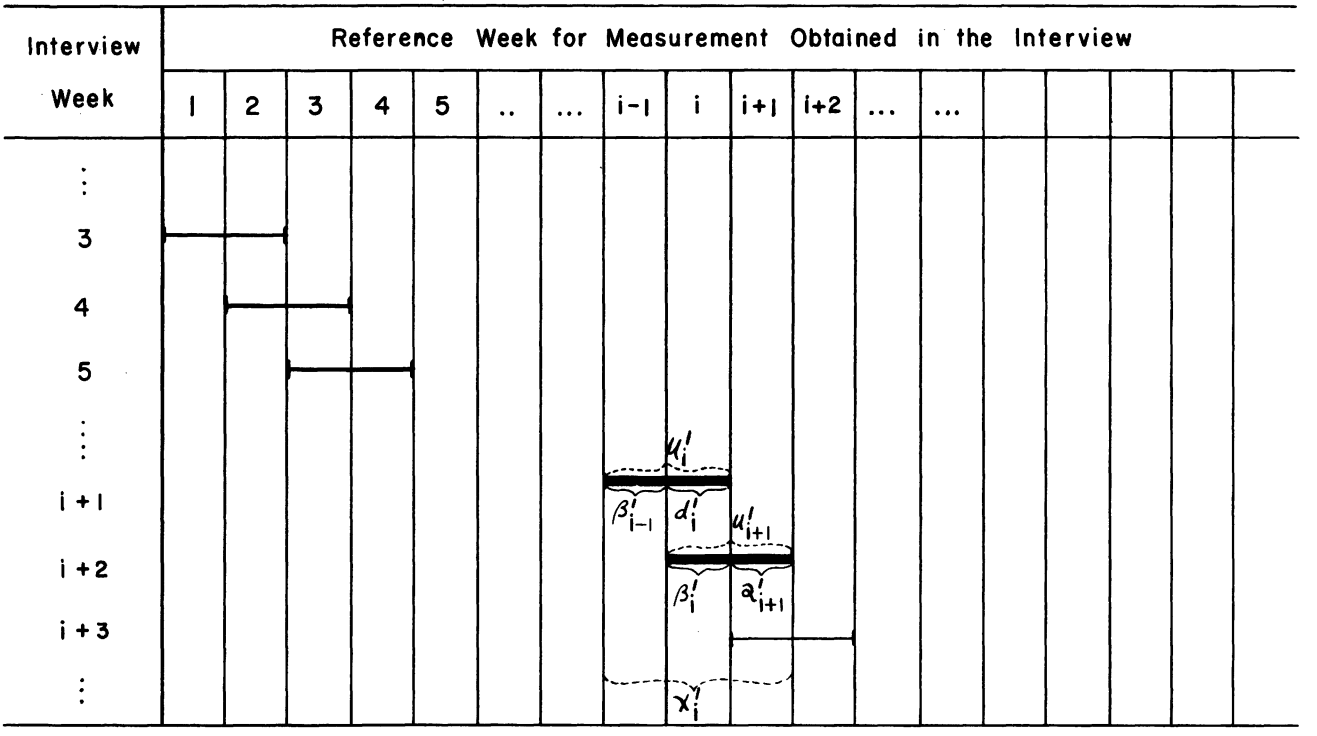
Other Technical Developments

During the past year most technical statistical work outside that connected immediately with the continuing household interviews has been related to testing or exploring certain aspects of possible alternative schemes of measuring health phenomena. This work is being done largely by other organizations under specifications and contracts prepared by the Health Survey. It includes such activities as:

A. Efforts to develop an effective standard questionnaire for taking medical histories as part of a procedure for a health examination survey, i.e., a survey in which medical and dental examinations are administered to a sample of persons. This work is being done by the Survey Research Center, University of Michigan.
B. Analysis of differences between information from a household interview and medical records on diagnoses of chronic disease, and on the fact of hospitalization, for a selected group of members of the Health Insurance Plan of Greater New York City.
C. Pilot study of hospitalization of decedents, since the household interview covers the living population only. This investigation is being conducted by the National Office of Vital Statistics.
D. Study of attitudes toward participation in a health examination survey, being conducted by the National Opinion Research Center.

These projects and others are currently in progress. A preliminary report on the attitude study

Figure 1. Relationship Between Interview and Reference Week for Selected Questions in the Health Household-Interview Survey. *And* Pattern of Estimates d_i' β_i' u_i' & χ_i'



The estimate $u_i' = \frac{1}{2}(d_i' + \beta_{i-1}')$

The estimate $\chi_i' = \frac{1}{2}(u_i' + u_{i+1}')$

just mentioned has been given by Borsky and Sagen[12] in a paper presented at the American Public Health Association meeting in St. Louis. The National Health Survey expects to institute a health examination survey at a later date. This was the reason for the study on motivation, since both evidence and opinion have suggested that nonresponse will be a major problem for the proposed health examinations. Borsky and Sagen concluded tentatively that favorable response is correlated with favorable attitudes toward doctors, medical research, and the government's role in health matters. They noted too, a close relationship between acceptance of the examination and the individual's concern over his own health. The next step is to see if these findings can be used in reducing the nonresponse rate.

The Health Examination Survey

In point of time the continuing household survey is the first major NHS activity. The second major NHS undertaking is expected to be a Health Examination Survey. As has been evident from previous remarks, this survey is in its planning and testing stages. The precise form which it will take has not been finally determined. The general plan contemplates the selection of a small probability sample of the population, and the administering of a limited single visit medical and dental examination to persons selected. The objective is to obtain medical statistics that cannot be gathered by other types of surveys, and in particular, to complement information obtained in the household survey.

Theoretical and experimental work in this survey have included in the past year—in addition to those items already mentioned—such matters as general survey and sample designs, proper content of examination, methods of conducting examination, and a pilot study of the project conducted in the Washington, D. C., metropolitan area. The pilot study applied to a sample of 180 persons, ages 18-64, who had been given the usual household interview and who had been offered the health examination. Examinations were actually given to 119 of the 180 persons. Experience in this pilot study is being analyzed while two more pilot studies are being planned, one in a smaller city, and one in a rural area. It is hoped that these

three studies will provide sufficient information that they may be followed by initiation of the Health Examination Survey itself.

Evolutionary Pattern of the National Health Survey

One final item should be added to this sampling of developments in the National Health Survey. This is an idea which was present at the start of the program and which has become one of the guiding principles of the Survey. It is that the National Health Survey is no single specific procedure, but rather is a coordinated set of undertakings intended collectively to provide intelligence on the health conditions of the population—and further that the Survey is an evolutionary program which must develop in response to need and experience.

REFERENCES

1. U. S. National Health Survey. Origin and Program of the U. S. National Health Survey. Health Statistics. Series A-1. PHS Publication No. 584-A1. Public Health Service. Washington, D. C., May 1958.

2. Public Law 652 - 84th Congress, Chapter 510-2d Session - S. 3076.

3. One aspect of this testing, concerning a pilot study of household interviewing, was reported on at the 1957 annual meeting of the American Statistical Association, by Harold Nisselson and Theodore D. Woolsey.

4. U. S. National Health Survey. Preliminary Report on Volume of Physician Visits, United States, July-September 1957. Health Statistics. Series B-1. PHS Publication No. 584-B1. Public Health Service. Washington, D. C., February 1958.

REFERENCES—Continued

5. U. S. National Health Survey. Preliminary Report on Volume of Dental Care, United States, July-September 1957. Health Statistics. Series B-2. PHS Publication No. 584-B2. Public Health Service. Washington, D. C., March 1958.

6. U. S. National Health Survey. Preliminary Report on Number of Persons Injured, United States, July-December 1957. Health Statistics. Series B-3. PHS Publication No. 584-B3. Public Health Service. Washington, D. C., May 1958.

7. U. S. National Health Survey. Preliminary Report on Disability, United States, July-September 1957. Health Statistics. Series B-4. PHS Publication No. 584-B4. Public Health Service. Washington, D. C., June 1958.

8. U. S. National Health Survey. Selected Survey Topics, United States, July 1957-June 1958. Health Statistics. Series B-5. PHS Publication No. 584-B5. Public Health Service. Washington, D. C., November 1958.

9. U. S. National Health Survey. The Statistical Design of the Health Household-Interview Survey. Health Statistics. Series A-2. PHS Publication No. 584-A2. Public Health Service. Washington, D. C., July 1958.

10. U. S. National Health Survey. Concepts and Definitions in the Health Household-Interview Survey. Health Statistics. Series A-3. PHS Publication No. 584-A3. Public Health Service. Washington, D. C., September 1958.

11. Provisional Tabulation From the U. S. National Health Survey. Current Statistics on Respiratory Diseases. Numbers 1-29, November 7, 1957-May 22, 1958.

12. Borsky, P. N., (NORC) and Sagen, O. K. (NHS): Motivations Toward Public Health Examinations. Paper presented at the American Public Health Association, St. Louis, Mo., October 1958.