WORLD HEALTH ORGANIZATION

REGIONAL OFFICE FOR THE EASTERN MEDITERRANEAN

ORGANISATION MONDIALE DE LA SANTE

BUREAU RÉGIONAL DE LA MÉDITERRANÉE ORIENTALE

SEMINAR ON VITAL AND HEALTH STATISTICS

Damascus, 19 - 27 October 1963

Agenda Item 8(e)

EM/SEM.VHS/16
15 September 1963
ORIGINAL: ENGLISH

SOME ASPECTS OF METHODOLOGY OF HEALTH AND MORBIDITY SURVEYS

by

H. Christensen
Epidemiological Studies Unit, WHO

Introduction

Survey methods are being increasingly employed for the purpose of obtaining information on the health status of populations, conditions influencing health and the role played by existing health services. Although many health surveys have already been conducted, especially in the highly developed countries, no standard survey methods or principles have been formulated or recommended for different kinds of objectives. In some respects standardization may be achieved, that is, particular methods may be accepted for universal application, whereas in others it may be unrealistic to lay down rigid rules to be followed under all conditions. Still it may be possible in the latter case to get to know the effect on measurements of the use of different methods so that it would be possible all the same to compare survey results based on different approaches. Comparative studies on existing and new methods of gathering information may provide valuable data to help formulate standardized procedures.

The present paper intends to provide a short review of the methods in use today in health and morbidity surveys where the source of information is interrogation or examination of individuals. It does not pretend to cover more than a limited aspect of the subject, but it may give an idea of the issues involved and stimulate discussion on the subject, the importance of which should be more fully appreciated.

Methods of Investigation

Interview surveys

In a health survey intended to provide information about levels of general morbidity, the utilization of health services and the cost to the population of maintaining health, etc., the method of collecting the desired information is normally a special interrogation of individuals carried out by means of household visits or interviews of single individuals, by trained interviewers. In this way it should be possible to measure the number of persons ill, the number of spells of illnesses, the duration of these illnesses, and also the amount and type of medical care sought and received, and the cost to the individual of promoting and restoring health, etc.

One of the limitations of this method of investigation as regards measurement of morbidity is that the diagnosis is of varying reliability; the use of lay investigators limits the precision of the questions that can be put as to cause of illness, and there may also be difficulties because of the person's ignorance of the nature of his "illness". Within these limitations, however, the interview method may nevertheless provide a picture of the extent and distribution of disease, including those minor ailments and disabilities which may constitute a fair proportion of the total illness of the population but which otherwise would not come to the notice of a health administrator.

The limitations referred to above should not however be underrated; the substitution of questions for medical diagnosis may not prove satisfactory in all situations and the outcome of the survey relating to the health status of the population might be that information on illness would only be available in broad and rather vague categories such as "fever", "stomach troubles", "other diseases not diagnosed", etc. The survey methods are closely linked with the survey content and the question arises whether only such surveys should be promoted as have a reasonable chance of success under the particular circumstances. It may be argued that the best possible information that can be obtained is better than nothing; this is true only to the extent that the information obtained is not misleading - that is, certain minimum requirements must be met as regards the validity and reliability of the data.

Considering the situation where it is desirable and feasible to carry out a health survey by means of interviews, a number of different approaches to collecting the information are possible. For instance, the interview unit may be the household or the individual; a single visit or more visits may be made; the survey period may be one day, one month, one or more years - or the survey may be continuous; probe questions may be used; diaries can be kept by individuals; the length of the recall period may be one day, one or more weeks, one or more months, or as much as one year in certain instances, as when information on hospitalization is sought.

In a household survey the quality of the information obtained depends in part upon the person who is interviewed. The best respondent for infants is presumably the parents, perhaps the mother. For adults, the interview of each family member would undoubtedly provide the most accurate results, but an approach of this type would require repeated visits to households in order to ensure a reasonable coverage. With a view to limiting the costs of the survey, it may be considered acceptable to compromise with this procedure and make use of proxy respondents for adults. This method is for instance applied in the United States National Health Survey where a responsible adult member of the household is questioned if a household member over the age of 18 is not present at the time of the interview. A different outcome in survey results may be expected when proxy respondents are accepted. In the pre-tests carried out in the United States National Health Survey a study was performed to evaluate the effect of the use of proxy respondents on morbidity data obtained, and it was found that this method not only reduced the total amount of illness reported but also resulted in a qualitative change of the composition of illnesses which were reported - the minor non-chronic conditions being relatively more under-reported. On the other hand, the average number of visits required per completed household interview was found to be considerably less when proxy respondents were accepted, thus reducing the field cost appreciably.

The use of proxy respondents is equally relevant in surveys where the individual is the interview unit. In the survey of sickness carried out in the United Kingdom during 1943 to 1952, a proxy closely related to the respondent was found whenever possible, after three unfruitful calls had been made. For those who could not be contacted, or for whom no "first-class" proxy could be found, a substitute of the same sex and age as the original person, and as near to him geographically as possible, was selected from among the extra names drawn for this purpose when the sample was taken. In the Sickness Survey of Denmark 1951/1954, where the individual was also the interview unit, no replacement of those who could not be contacted (after four calls) was made, and even relatives in the house were not allowed to answer the questions although they were permitted to give assistance. This implied that the actual number interviewed fell short of the sample population envisaged by 20 per cent.

As regards the number of calls made on the interview unit, some public health administrators have favoured the method of questioning the selected persons of households only once. In this way a relatively large part of the population could be covered; further, all persons interviewed are equally unprepared when answering the questions, whereas in the follow-up system there is a risk that the respondents will give more, and unnecessarily detailed, information during the latter part of the survey than when first questioned. The single interview survey on the other hand also has its drawbacks. For instance, information on duration of illness may be less accurate than when repeated visits are made.

In order to obtain information on seasonal variations in the extent and distribution of illnesses, injuries and other topics, the duration of a time-limited survey should not be less than one year. The information may be obtained by following the same sample over a full year as was done in Japan in 1951/1952. However, more reliable and more easily obtained information on seasonal variations may be gathered by repeated surveys over the year with a different sample population each time. The survey carried out in Japan mentioned above, demonstrated that repeated visits to the same household resulted in declining interest and enthusiasm on the part of the population as time went by. It may thus be generally concluded that in lengthy and continuous surveys the use of the same sample population is not advisable.

A variation of a continuous survey is the periodic survey - designed to keep morbidity and the medical care system under regular review. An example of this type of survey is the National Health Survey in Japan which is conducted during one month every year - usually in November. However, such a one-month survey does not give information on seasonal variation of the extent and distribution of diseases.

Also to be considered is the length of round - that is the period for completing the interviews of all households or persons in the sample. Just as there may be a practical limit to how fast a round can be completed, there would also be a limit to how long it would last. For instance, if information on morbidity is sought for a certain month and a round is scheduled to last one month, the first person questioned

would have to try to recall what happened during the previous month, whereas the last person questioned would have to remember two months back and ignore events which occurred during the preceding month. This would most likely lead to difficulties in getting an exact monthly division of incidence, and further, a deficit in prevalence rates might result because more recent events are better remembered than earlier ones. It would therefore be advisable to complete a round fairly quickly.

A different approach would be to study a period which ends on the day of interview; for instance the informants may be asked about their health during the last two weeks. This would exclude the possibility of the erroneous inclusion of events which occurred after the study period but before the interview, and would also ensure that all informants had to retrospect equally long. One disadvantage is that people will be questioned about rather different periods of time, in particular if a round is lengthy, and the importance of this consideration depends on the seasonal variation of the factors studied.

The memory of the respondents may be refreshed by asking a variety of questions which would also ensure that they understand the range of physical and medical defects which the survey is intended to cover. It is however important that the interviewer does not suggest an answer or cause the respondent to produce an incorrect reply. In some surveys a check-list is used in the final phase of the interview; for instance it may be asked: "to make sure, may I ask you whether you have had anything wrong with your eyes, ears, stomach, etc., during the month?" - thus mentioning distinctly the various organs according to a pre-designed list. In this case great care must be exercised not to emphasize some entity more than others. It should be realized that the use of check-list questions would measure population morbidity at a higher level than when no check-list is applied, and that it is therefore essential that the interview design is standardized and uniformly applied throughout the survey.

One method of overcoming some of the problems of the memory factor is to use diaries which all the informants should be asked to maintain every day during the survey period. Such diaries are an aid that can be utilized only when repeat visits are made.

The memory of informants is a major factor in deciding the length of recall period to be adopted in the survey, the recall period being the period of time to which questions on health and other topics relate. A very short recall period, say the last 24 hours before interview, would for instance give rise to a smaller number of recorded illnesses than when a recall period of, say, two months is in question. In the latter case the data collected would be less reliable than in the former because of memory loss; on the other hand, for a short recall period a larger sample would be required in order to obtain sufficient data for analysis. Also the proportion of each disease to the total may come out differently for a different length of recall period. The influence of the memory factor was studied in the Californian Health Survey carried out in 1954/1955. By comparing results relating to recall periods of from one to four weeks, it was found that the memory loss increased with the increasing length of the recall period, and that minor diseases or ailments were influenced by the memory loss more markedly than more severe ones. Short recall periods have been used in the Canadian Sickness Survey 1950/1951, and in the periodic sickness survey in Japan, but in both instances repeat visits and use of diaries have been made. A recall period as long as two months was adopted in the United Kingdom Sickness Survey.

Aside from various possibilities of interview design and different ways of dealing with factors relating to respondents, there are also various methods of training and utilizing interviewers. As the order and form of questions is decisive for the results obtained, the way in which the interviewers put their questions (if any degree of freedom is allowed) and record the answers, influences the results. Lay interviewers thoroughly instructed in disease terminology might be tempted to interpret the respondents' answers and biased results may ensue. In training, emphasis should be put on the importance of an unbiased interview and it appears a good rule that the information should be noted down immediately it is given, as was done in the United Kingdom Sickness Survey. As important as suitable training for the success of a health survey is the problem of ensuring a suitable interviewer work load. The number of interviewers employed would depend on the average length of each interview and the total time allotted for a complete round. This and other matters of an administrative character can best be investigated in a pilot phase.

The above is a short and far from complete review of certain methodological points regarding the collection of information in interview health surveys. A number of countries have already carried out interview health surveys but they all differ as regards methodology, quite apart from differences in survey content. As different methodology clearly gives rise to a different outcome in results, it is impossible to compare directly the findings of these surveys. The same would apply to surveys within a country if different methods were employed. Further, a change of methodology in the course of a survey might well create serious problems in the interpretation of results. In deciding on the methodology to be adopted, the problem of comparability must be carefully considered, not only with regard to repeated surveys in the same area but possibly also with a view to comparisons with surveys in other areas or countries. Before embarking on an interview health survey it would be useful to study methodology employed elsewhere and profit by the experience obtained. Survey reports should, when published, include a detailed description of the methodology used for a correct appreciation of the results and also for the information of others interested in health surveys.

Health examination surveys

When it is important to obtain precise diagnostic information, the survey must obviously include clinical examination and appropriate diagnostic and laboratory tests. The diagnostic procedures should be fairly simple, however, as it is difficult to carry out elaborate examinations in the field. Whenever possible, it is desirable to use screening techniques for selecting persons for more thorough examination.

The advantages of the health examination method are: (a) improved accuracy of diagnosis as compared with the interview method, and (b) as in the case of interview surveys, provision of a population base on which meaningful morbidity rates can be calculated. A limitation of this method of investigation is the cost involved, even when very simple diagnostic screening procedures are used. Surveys based on special medical examinations are therefore suitable principally for estimating the prevalence of specific diseases in a fairly small area so as to help formulation of effective control programmes.

In planning a health examination survey it should be ascertained that simple methods exist for giving valid results - that is, the methods used should be sufficiently specific and objective to be suitable for the purpose. A clinical examination or diagnostic test should therefore measure with reasonable accuracy what it purports to measure. Evaluation of validity may be made in pilot studies and again built-in assessment should be considered when possible by performing other measurements known to be associated with the item under investigation. For example, in surveys of chronic respiratory disease in adult men, the ventilatory capacity measured by a peak-flow meter would be expected to be highly correlated with the doctor's diagnosis of the condition and with the answer given to standardized clinical examination.

However, a valid method of investigation does not necessarily of itself assure precise information because variations may arise in examination techniques as well as in the interpretation of the findings. Before adopting a method of investigation it is important to know to what extent it can reproduce the findings, or in other words the degree of inter- and intra-observer variation. In this connexion it should be emphasized that observer variability can be substantially reduced by employing qualified investigators who have received appropriate training. Whenever possible each individual survey should be designed to include its own tests for assessing observer variability. As an example of this it can be mentioned that in tuberculosis surveys assisted by WHO, it has been customary to check regularly the consistency with which the same observer reads tuberculin reactions and also the variations in readings between different observers.

From the foregoing it will be seen that for a health examination survey, the diagnostic procedures should be sufficiently precise, and at the same time not expensive or difficult to apply in the field. How precise the diagnosis should be depends on the objectives of the survey and proposed future action. This is a subject to which health officers should give considerable thought.

Sampling

A survey may involve either complete enumeration of the population of a country or of a specified locality; or enumeration of samples only of the population. In most instances, complete enumeration in unnecessary, time-consuming and expensive. Nowadays sampling methods are being used increasingly and such methods form an essential part of survey methodology.

In sampling, a part of the population is examined, usually a small part, and based on these findings estimates are made for the entire population. In order to make valid estimates for the population the sampling has to be done on sound statistical principles which means that every member of the population has a pre-determined chance of being included in the sample. Unless this fundamental requirement is satisfied it is not possible to arrive at any prediction for the population as a whole. In general, simple sampling procedures ensure that every member of the population has the same chance of being selected for examination or interview, but certain refinements of method exist where greater efficiency may be secured by selecting different segments of the population with a varying chance of being included in the sample. Sampling theory and methods are highly developed, and several techniques are available to suit particular situations and ensure maximum efficiency. It may be pointed out at this stage that even with the most refined techniques a sample survey cannot be taken to be identical with a complete enumeration. Nevertheless, a statistically chosen sample enables prediction to be made on the population within certain limits and associate with it a certain probability. An example would make this clear. Suppose a sample of the population is examined for a particular morbid condition and it is found that six per cent. of the persons examined suffer from this condition. Depending on the sample size and the method of sampling, the prediction for the population would be somewhat as follows: the proportion of individuals with the morbid condition was found to be six per cent. in the sample and the true proportion in the population is most likely to be between five and seven per cent. (say), unless a chance of one in 20 had occurred. Of course it is possible occasionally to hit on a sample that gives an estimate far different from the true population value. For example, the true value in the population could be well outside the range mentioned earlier. However, the chance that we are going to be so unlucky is only five per cent and that is a risk worth taking.

The main advantage of carrying out a sample survey is that limited personnel, resources and facilities can be employed to carry out intensive examination. In most countries there is an acute shortage of qualified personnel and funds are limited. A complete enumeration carried out by available personnel would mean that efforts were scattered over a wide range, with the result that the actual examination procedures become less intense or even superficial. There is also another very important reason for choosing a sample. In any type of survey, whether of the complete enumeration type or of the sample survey type, there is bound to be a certain number of individuals who are either uncooperative or escape examination for various other reasons. These reasons may have an important bearing on the validity of the results. For example, in a leprosy survey many patients actually suffering from leprosy may elude examination. In a complete enumeration the number of such persons is likely to assume a serious proportion and it may be extremely difficult and costly to track them down. The same problem of non-response may arise in the case of a sample survey as well, but then the number of such non-respondents becomes proportionately small and intensive efforts could be made to trace and examine them.

In the above-mentioned hypothetical example it was mentioned that from the sample value for a characteristic it would be possible to say in what range the true value in the population lies. The question immediately arises how wide the range should be. Is it enough if the prevalence of a disease in a community is estimated to be between five and seven per cent. or would it be more desirable to say that the true prevalence should lie between 5.5 and 6.5 per cent? In other words, how "precise" is the estimate required to be? This immediately raises the question of sample size. Here it must be mentioned that if double precision is desired the sample size should be quadrupled. To determine the sample size for the survey of any disease, two basic items of information must be furnished, viz., approximate prevalence of the disease and the required precision of the estimate. The first of these items may seem absurd because the very reason for conducting the survey is to get an idea of the prevalence, but in practice, based on past experience, health workers do have an idea in what range of magnitude the prevalence should lie. For example, there may be reason to believe that the prevalence of, say, pulmonary tuberculosis is in the

range of say two to four per cent. In such cases the lower estimate should be chosen in order to arrive at the sample size, for the smaller the prevalence the greater the sample size necessary to be able to estimate with the same degree of precision. A balance has to be struck between desirable precision and cost. In morbidity surveys it is meaningless to ask for a very high degree of precision when the errors of measurement due to the diagnostic tools employed are very high. It would be much more profitable to use the same money to improve diagnostic techniques rather than ask for a high degree of precision.

There are various sampling techniques available, and the choice of a suitable technique in a particular situation depends on a variety of factors, e.g. practical considerations, how detailed the information should be, the precision required, etc. For the same cost the technique of choice would be the one which gives the maximum precision. A very brief review of some of the commoner techniques may not be out of place at this juncture.

Simple random sampling

This is the simplest of the techniques and in fact is the basis of almost all the other techniques. Here individuals or groups of individuals, e.g. households, are chosen at random from the whole population, as, for instance, from a list, until the required size of the sample is obtained. The random selection is generally made by numbering the population and making use of a table of random numbers to choose members of the sample.

Stratified sampling

In this type of sampling the population is first divided into groups or "strata", each stratum being sampled separately so that a specified number of sample units is obtained from each stratum. Such strata may be geographic sub-divisions, divisions depending on occupation, industry, etc. The technique of stratification, when the population is subdivided into groups or strata in such a way that units in each stratum are as similar as possible, might substantially reduce the sampling error. The sampling rate (i.e. the proportion of units that are selected) may be varied from one stratum to another, with the possibility of a further increase in precision for the same sample size.

Multi-stage sampling

In this sampling, the population is regarded as made up of a number of first-stage units each of which is composed of second-stage units, and so on. For example, a country is divided into a number of districts, each district into a number of still smaller sub-divisions and so on. A number of districts is selected in the first stage; within each selected district a number of smaller sub-divisions is selected in the second stage, and from each selected sub-division a number of households may be selected at the third stage. Sometimes stratification is combined with multi-stage sampling at any particular stage or at all stages.

Multi-phase sampling

In this procedure certain items of information are collected on all the units of the sample and other items of information are collected on a sub-sample of the latter. This is called a two-phase sampling procedure. This could be extended to three or more phases. As an important application, the information obtained in one phase could be utilized to provide more accurate estimates of the variables obtained at a later phase.

Interpenetrating samples

This is an efficient device to estimate non-sampling errors such as observer bias, etc. When the survey work is being carried out by several teams, independent samples may be drawn so that each sample may be investigated by a different survey team. Some of the units may be kept identical so that observations on the same set of individuals by two separate teams may be compared.

An attempt has been made to describe very briefly what some of the commoner sampling techniques imply. In practice sampling surveys are usually combinations of the techniques to suit particular local conditions.

It may be useful to draw attention to certain terms used in sampling which are being used frequently. The United Nations Statistical Office have defined some of these terms which are reproduced below.

Frame. The frame or "substrate" consists of previously available descriptions of the material in the form of maps, lists, directories, etc., from which sample units may be constructed and a set of units selected.

Elementary units. Units which are the smallest parts of the material capable of possessing a particular characteristic. Elementary units may, therefore, be of different types (e.g. number of children is a characteristic of a family, age is a characteristic of an individual), for example, at different stages of multi-stage sampling.

Sample units. The units which form the basis of the sampling process. The sample units may be (a) the same as the elementary units, or (b) groups of such units capable of possessing all the characteristics required. A group may consist of a cluster of contiguous elementary units, or a number of units arranged in an assigned configuration. A systematic pattern of elementary units may, for instance, constitute a sample unit.

The frame would define the geographical scope of the survey and the categories of material covered. For example, the frame could be a census list, in which case it should specify what areas the census has covered, what categories of people have been included, and the date and time of the census operation. If for some reason it is decided to exclude certain areas of the country, the frame would be the total area less all the excluded areas. Similarly, if for example in Africa it was decided to draw a sample of Africans, the frame would consist of African nationals only.

As examples of sampling designs of health surveys the procedures employed in the current United States National Health Survey and in the Canadian Sickness Survey, 1950/1951, are briefly described below.

The recent health survey in the United States of America adopted a multi-stage design introducing stratification. The United States was divided into an artificially-constructed 1900 Primary Sampling Units (PSU's), based on certain geographical and administrative considerations as also on the available information on descriptive statistics. The population of each PSU was at least 7500 in the Western United States and at least 10,000 in other parts. The PSU's were stratified

in such a way that members within each strata were more or less alike. The principal modes of stratification were geographical location, density of population, rate of population growth between the last two population censuses, proportion of non-white, type of industry in predominantly urban areas, and type of farming in rural areas. There were altogether 372 strata, and one PSU was selected from each stratum for inclusion in the sample with probability proportional to size. Thus 374 first-stage units resulted. For sampling purposes, and in order to reduce the over-all variance, the civilian population of the United States was divided into three mutually-exclusive classes or selected zones:

Zone A: those persons living in common dwelling places

Zone B: those persons living in areas of "new housing"

Zone C: those persons living in large special dwelling places

Later-stage sampling was handled separately for each zone, making use of the National Housing Inventory of 1956. The details of selection may be found in the report of the United States National Health Survey (the statistical design of the health household-interview survey) Series A-2.

The ultimate sampling unit within the PSU is called a segment. It is a geographically-defined parcel which contains an expected six households. Segments to be included in the sample are chosen in a series of steps or stages. For each week a random sample of about 120 segments is drawn. The household members interviewed each week are an independent representative sample of the population, so that the samples for successive weeks can be combined into larger samples for, say, a three-month period or a year. The national sample plan over a twelve-month period includes approximately 115 000 persons from some 36 000 households in about 6000 segments.

In the Canadian Sickness Survey in 1950/1951 (covering approximately 10 000 householders distributed throughout the 10 provinces in metropolitan, small urban and rural areas) the population data used for selecting the sample were drawn from the 1951 census of Canada, except in the case of Newfoundland where a census taken in 1945 supplied the necessary information. The basic unit of selection was the dwelling, defined as "a structurally separate set of living premises having its own

entrance from outside or from a common passage or stairway inside". All the 28 metropolitan areas with a population of at least 30 000 were included in the sample, and the chosen dwellings for these centres were selected in two stages: a sample was chosen from lists of city blocks and peripheral segments, and from these lists a random sample of households was selected. The remaining area of each province was divided into "primary sampling units", consisting of single or grouped municipalities or census sub-divisions, each with a population of about 2000 persons. The primary sampling units were then stratified by allocating each of them to a group or stratum on the basis of contiguity and similarity of certain population characteristics A number of strata was thus established in each province or region, each consisting of approximately 30 primary sampling units or about 60 000 persons. The next step was to divide each stratum into two substrata, one containing the urban primary sampling units and the other the rural primary sampling units.

Urban substrata

Within each urban substrata a single primary sampling unit was chosen by probability sampling, each unit's chance of selection being proportional to its population size within the total population of the substratum. The actual selection of dwellings depended primarily on the availability of reliable directories, tax assessment rolls, or similar lists of households. When these were prepared, a systematic sample was drawn from the numbered list by taking a random start and then applying the denominator of the sampling ratio as an arithmetical progression down the list. The sampling ratio itself was obtained by applying the predetermined provincia sampling ratio to the population of the substratum and dividing the resulting figure by the population of the selected primary sampling unit.

Rural substrata

Each rural substratum was further subdivided into two smaller groups of primary sampling units on the basis of high or low average farm income. From each of these resulting smaller substrata a random selection of a single primary sampling unit was made. The sampling ratio for each was calculated in the same way as for urban primary sampling units.

In many cases, the rural primary sampling units which were drawn into the sample were quite large, with dwellings thinly scattered throughout, so that a systematic sample drawn directly would have confronted enumerators with an extensive amount of travel. In order to reduce the amount and expense of travel, the households were first aggregated or "clustered" in groups of three or multiples of three by the use of large-scale maps showing households as dots and indicating natural boundaries, such as rivers, roads and railway tracks, easily recognizable in the field. The clusters were then sampled systematically, and the households in each single cluster were listed for enumeration.

Questionnaires and Record Forms

The framing of questionnaires to be used in the course of the survey is one of the most important tasks at the planning stage. It is difficult to lay down detailed rules as to what questions should be asked and how, but certain general principles are common to all types of questionnaire.

First of all, the investigators must have a precise idea of what they are going to do with the information collected. How often are questions included because someone has a vague idea that they may be interesting? The actual procedure in practice would mean the listing of all possible items of information that may be considered relevant in this connexion, carefully eschewing items that are redundant. The list of items may be further shortened as usually it cannot be undertaken to obtain answers to all the questions which can be considered of interest.

Having selected the topics to be included in the inquiry, the questionnaire or record form can be drafted. The construction of statistical schedules is something which is best learned by actually making and using them. Nevertheless, certain points may be mentioned. The form or questionnaire should be so devised as to elicit the information in a logical and cogent way and should permit easy recording. The entire form, as well as each question, should be simple and as clear as possible. The ambiguous question or the question that invites an ambiguous answer produces useless data and involves a waste of time and money. A simple example of this is asking the question: "married or single"? A meaningless answer of "yes" or "no" may result, and further, not all persons are included in these two categories. For instance, "single" may cover the situations of "never married" or "divorced".

Certain types of questions should be omitted. For example, leading questions should be scrupulously avoided. Questions which are unduly inquisitive, or which are likely to give offence, should likewise be avoided. Questions concerning personal matters (such as income) should be handled with tact and it may be the best approach to deal with such matters at the end of the interview after the co-operation of the informant has been obtained.

Answers should be objective and capable of tabulation. Subjective terms like "good" and "bad", etc. are very vague and not comparable. For example, asking an enumerator to record whether the ventilator in a house was good or bad would be very loose and not permit comparison between different enumerators. It would be preferable to ask the enumerators to record the sizes of the windows, etc., in relation to total area, so that the medical officer in charge could later classify the quality and nature of ventilation according to an objective criterion. Although physical measurements are more objective, qualitative observations may, however, often be capable of summing up the salient features of a complex situation. For instance, the qualitative grading of dampness of a house is likely to be more effective than the physical measurement designed to determine the degree of dampness. Qualitative observations can themselves be made objective when proper standardization is ensured.

Instructions and definitions should be concise. The enumerator and informant should never be in doubt as to what information is desired and what terms or units are to be used. When inquiring as to the health status of an individual, the inquiry must refer to some specific time. It should be clear what is meant by "household", whether age is to be "to nearest birthday" or "to last birthday", what categories are to be used for "race", "occupation" and "industry", etc. If information is desired as to number of rooms, it should be noted whether or not bathrooms, kitchens, dressing rooms and the like are to be counted as rooms.

The arrangement of questions should be carefully planned. The order of the questions should be such as to facilitate the answering of each question in turn, and, what is very important, sufficient space should be allowed for recording.

As a rule, the form should be designed in such a way as to facilitate mechanical punching and sorting. Even if hand sorting is resorted to, this method of design has been found, in practice, to be helpful. Careful consideration has to be given to sorting procedures. Sometimes when a mass of valuable data has been collected it is found that cross tabulations bringing out the correlation of several characteristics are not easy because not enough thought has been bestowed initially on the design. It is advisable to think of all possible ways of analysing the data beforehand, and in fact make all the specimen tables that will be used in the final analysis.

The schedule should be supplemented by a detailed manual of instructions, which must foresee all possible answers to a question and lay down rules as to how different types of answers are to be recorded. In surveys which involve observation and measurement, a separate set of instructions is usually prepared and the form would not carry its own full explanation. With the general form of question, i.e. when the question is not put in the exact form given, it may often be preferable to give explanatory notes on the form itself, so as to ensure that the investigator knows exactly what information is required.

In all cases it would be advisable to test the questionnaire or record form in a pilot run on a small section (not necessarily a random sample) of the population, discover its shortcomings, and then revise it in the light of the try-out. Experience has shown that such pilot studies have saved a great deal of labour and trouble which would otherwise have resulted.

Control of the Accuracy of Work

Plans for the control of the quality and accuracy of the information collected are essential parts of the methodology of surveys. Carefully designed questionnaires and refined sampling schemes could be rendered fruitless if sufficient attention is not devoted to this aspect of survey work. Supervisors should exercise field checks on a random sub-sample of units which should be conducted in such a manner that the investigators cannot know which parts of their work will be checked. Checks of this type are rather difficult in the case of the interview type of survey. However, in the United States Health Household-Interview Survey the re-interview procedure was

used in connexion with training and supervision of interviewers. A sixth of the households assigned to each interviewer was re-contacted by his supervisor regularly and one pre-designated member of the household was interviewed. The objects of the re-interview procedure were threefold: (a) the training and quality control of the interviewers, (b) the measurement of interviewer variability, and (c) detection of interviewer bias. The use of interpenetrating samples described earlier under "types of sampling" could also serve as a check against major defects in individual investigators.

The examination of the completed forms should be a regular feature of the survey. Preliminary examination will reveal major shortcomings and defects in the form itself. If the number of returns is large and the interviewers or enumerators are properly trained, it may be sufficient to examine a random sample of the completed forms thoroughly.

The questionnaire should generally be so designed that completion on the spot is possible. If, however, it is necessary for the interviewer to take notes and complete the form later, such completion should be done as soon as possible. As a rule each form should be completed neatly as and when the interview is being held. The principle of "fair copying" from a "rough" original introduces another and unnecessary source of error - viz. copying errors. If such "fair copies" are made the originals should be sent along so that a sample check can be made to control copying errors.

The follow-up of non-respondents is another part of a survey programme that has to be given careful thought. In a sample survey the substitution of one household for another is absolutely forbidden as it will completely upset all estimation procedures. Every effort must be made to follow-up non-respondents to find out the causes of non-response as they may have a vital bearing on the estimates. If the proportion of non-respondents is large, a sub-sample of them at least should be followed-up. In general some very simple sampling methods, such as taking of every q.th. non-respondent is adequate.

Control of accuracy does not stop with control of the enumeration technique. A good deal of editorial work remains to be done in the central office. The completed questionnaires have to be carefully examined to detect gross errors and editing generally. Examples of gross errors that could creep in are that a male has borne children or that a child of five is an engineer. Editing consists broadly of checking for consistency, detection of gross errors and interpreting illegible writing and answers in a form that can be coded.

In health surveys coding of the information on causes of sickness is an expertise job. The enumerators will generally be required to record the causes of illness just as reported. It is advisable to have these causes coded according to an established classification, such as the International Classification of Diseases, at the headquarters of the survey organization. It would be useful if two or more coders could be assigned to do the coding independently, any differences to be decided upon by an expert.

Processing of Survey Data

Processing refers to the methods of handling the data. The data have to be extracted from the questionnaires in predetermined tabular form either for summarizing the information on the community or for bringing out relationships between different characteristics measured. There are several ways of doing this. One could, for example, make all types of analysis directly from the cards. However, this possibility has to be ruled out except in the case of surveys where at the most a couple of hundred cards are involved. In small surveys involving say 500 cards or so, Cope-Chat cards could profitably be used. These are cards with holes round the edges. The holes can be made to correspond to answers to specific questions. For example, one could have 10 holes to represent 10 age-groups. By means of a punch, V-shaped notches can be cut out of the card. If the cards are arranged in a pack and a needle is passed through a particular hole, the cards punched in this hole will fall from the pack when the pack is lifted. Thus cards can be sorted into different classes.

However, in large-scale health surveys, even the Cope-Chat cards would be inconvenient and cumbersome to handle. The use of punched cards is the most convenient. There are several systems available, the IBM, Hollerith, Powers-Samas, Bull, etc. Essentially the principle is the same. For example, a normal Hollerith or IBM card contains 80 columns, each column containing digits numbered from 0 to 9. Each item of information collected on the questionnaire can be coded in such a way that it can be made to correspond to a particular column and digit in the punch card. Column 6 for instance could be made to represent sex and "1" in that column "males" and "2" "females". If one of the respondents in the survey is a female, a hole would be punched on "2" of column 6. Modern electrical punching devices would facilitate high speed punching operations. All the answers on a questionnaire would be transformed into holes on a punch card. The punching is verified on a verifying machine which would pick out cards incorrectly punched. The cards could then be fed into a sorter for sorting according to any desired classification scheme. If the survey is being conducted on a continuous basis it may be worthwhile to instal a machine in the central office of the survey organization. For ad hoc surveys it may be more economical to have the punching and sorting done in another office where mechanical equipment has already been installed, or to have the work done on a service basis in an office, such as the IBM dealing with such equipment

Nowadays there are various types of mechanical equipment. It is outside the scope of this paper to go into the details of the several equipments and processing techniques. It may, however, be mentioned that there are two broad categories: the conventional equipment and the electronic data-processing equipment. The conventional equipment uses the punch card, and they could range from a simple hand punch and the ordinary sorter to electrical punching machines and advanced sorting, counting, printing and tabulating equipment where several classifications may be accomplished by feeding the punched cards once into the machines. To these could be added the calculating punch which could perform most of the ordinary calculations needed in statistical analysis. In electronic data processing the information on the questionnaire is coded and by various means put on a magnetic tape. The tape is then

fed onto a computer which performs most of the calculations necessary with astonishing rapidity. Such equipment is, however, very expensive and requires trained personnel to prepare the "programmes" for the machines. It is generally more economical to share such equipment with other agencies which might have the possibility of using such equipment.

The final statistical analysis of the data depends on the specific purposes of the survey, what it is intended to find out or demonstrate, and on the design of the sample if a sample survey has been conducted. The main points in the analysis of sampling enquiries refer to estimation procedures for the parameters and measurement of their sampling errors. These jobs are often mathematical exercises and are best left to qualified statisticians.

REFERENCES

The Committee on the Danish National Morbidity Survey (1960), The Sickness Survey of Denmark 1951-1954, Copenhagen, Munksgoord

Health statistics from the United States National Health Survey: origin and program of the United States National Health Survey, Series A-1 and the statistical design of the Health Household-Interview Survey, Series A-2, United States Department of Health, Education and Welfare, Public Health Service

Logan, W. P. D. & Brooke, E. M. (1957) The survey of sickness 1943 to 1952, Studies on medical and population subjects, No. 12, H.M. Stationery Office, London

Nisselson, H. & Woolsey, T. D. (1959) Some problems of the household-interview design for the National Health Survey, J. Amer. Statist. Ass. 54, 69

Peart, A. F. W. (1952) Canada's sickness survey: review of methods, Canad. J. publ. Hlth, 43, 401

Statistical Office of the United Nations Statistical Papers Series C, No. 1 (revised), The preparation of sampling survey reports

Takemene, S. (1961) A health survey on Japan: family living studies, Symposium, International Labour Office, Geneva

Yates, F., Sampling methods for censuses and surveys, second ed., Charles Griffin & Co. Ltd., London