Margaret D. West and Garrie J. Losee, Bureau of Health Manpower

The Colombian National Health Survey was conducted by the Colombian Ministry of Health and the Colombian Association of Medical Faculties between September 1965 and June 1966. This survey was designed to provide information on health characteristics of the population of Colombia, a developing country of 17 million people. Data collected from 8,800 household interviews and 5,000 clinical examinations are now being analyzed and published.

Since 1957 the U.S. National Center for Health Statistics has been conducting a continuing Health Interview Survey. Beginning in 1959 a series of clinical examination surveys or cycles, directed toward various age-segments of the United States, has been conducted as components of the Health Examination Survey. The first of these cycles is singled out for comparative purposes, because it provides the widest coverage of the U.S. population to date. For the purpose of our presentation we will refer to these two data collection activities together as the U.S. National Health Survey, although, in fact, the National Health Survey is a broader program including other data collection mechanisms. We do not intend this comparison of methods of the two national health surveys to be an evaluative one. The purpose, rather, is to describe deviations in methodology in the Colombian survey from what could be called the "standard" methodology of the U. S. National Health Survey. We and the other architects of the Colombian health survey would have been satisfied to have been able to duplicate the highly developed methods of the U.S. National Health Survey. It should be mentioned that all but one of the seven United States consultants to the Colombian Survey were at the time directly associated with the U.S. survey The Colombian staff members of the survey observed the operations in the United States and consulted at length with their counterparts in the National Center for Health Statistics and the Bureau of the Census. It is therefore apparent that a definite effort was made to transfer knowledge and experience gained in the United States to Colombia.

However strong the desire for duplication may have been, it was apparent that modifications would have to be made in the "standard" methodology to accommodate both the special set of limiting conditions in Colombia and the different objectives. Major differences between the United States and Colombian health surveys are not many, nor did they occur by chance. We believe it is worthwhile to describe these differences, why they were felt to be necessary, and their implications in order to create a better understanding of the factors which can affect the design and operational procedures of a health survey, particularly one which is to be conducted in a developing country. In transferring, or adapting the U. S. methodology to Colombia we had to make adjustments to meet a number of special problems. At the same time there were advantages in the Colombian situation which added to the accomplishments and satisfactions of the undertaking.

Among the problems, five should have particular mention. The first of these was the limited experience and understanding of the concept of a probability sample. The only Colombian experience with sampling on a more than local basis of which we learned was the use of sampling techniques in agricultural studies. The 1964 National Census in Colombia was a 100% enumeration, with no sample for any aspect of the undertaking. In 1964 the National Statistics Department (DANE) considered requesting the assistance of a UN Expert to select a sample of the Census for advanced tabulation of the 1964 Population Census. Research in the social sciences was limited to non-probabilistic case studies. A second and related problem was the limited understanding of computer technology. This problem will be discussed later under the heading of data processing. Third, since the last Census had been taken in 1951, 13 years earlier, population counts, maps, and other sampling materials from that Census were considerably out of date. Special effort and methods were required to overcome these limitations in selecting an area sample. Fourth, the terrain of Colombia -- the several series of high mountain ranges which can be crossed only by airplane, the areas which can be reached only by incredible mountain roads, or only by boat or on horse back -- was a challenge both in the design and execution of the survey. Fifth, the high rate of illiteracy caused problems in the availability of interviewers and other staff. The illiteracy level also had to be taken into account in designing questionnaires and examination procedures. Added to these problems was the threat of political disorder. Concern had been expressed over the possible erruption of violence after the forthcoming 1966 national elections similar to that which followed the 1948 elections. Guerrilla warfare continuing from the 1948 disorders raised doubts about the feasibility of conducting a survey in some sections of the country.

Set against these problems were a number of real advantages. First was the sense of national purpose of the Colombians. This study had urgent meaning to them as a keystone of the development of a national health plan. It was undertaken with the active support of the President of the Republic and his Cabinet. seven medical schools of Colombia, through their Association, took active part in the planning. They provided medical students as interviewers, and residents from teaching hospitals as clinical examiners. These young people, in turn, brought to the study a sense of discovery and interest in their own country which was evident as enthusiasm throughout the survey. This, together with the excellent staff of the Ministry of Health, made for high professional standards for the work. Together, these assets made also for remarkably fine community cooperation, and in turn for very high response rates both for the household and the clinical surveys.

OBJECTIVES

The programs of the U. S. National Health Survey were developed on the basis of recommendations made in 1953 by the Subcommittee on National Morbidity Survey of the National Committee on Vital and Health Statistics. This group recommended a continuing national morbidity survey to provide "needed information on illness in the United States." It identified the potential uses of such survey data as:

- . Guide to administrative planning and evaluation of official and voluntary programs in the field of health.
- Evaluation of current morbidity experience in relation to the provision of services, facilities, and personnel for meeting the health needs of the nation or of a community.
- . Statistics for medical research.
- Statistics for drug and appliance manufacturers.
- . Statistics for public health education programs.

The committee recommended that a continuing survey be undertaken, and to meet the need for data on undiagnosed and nonmanifest disease, that there be physical examinations of a random subsample of the national sample.

The Colombian National Health Survey had its beginnings 10 years later, in 1963, when a group from the two Americas met in New York under the sponsorship of the Pan American Health Organization and the Milbank Memorial Fund. Their charge was to "design an appropriate research approach to the problems of physician needs and of medical education, to discuss the methodology to be used, and to define the appropriate emphasis and parameters of the studies."

This group recommended that one or two Latin

American countries be selected in which to undertake, on a pilot basis, a broad series of studies which would include:

- . A measurement of the health of the people and health services received in relation to demographic characteristics.
- . An inventory of health manpower resources.
- . A parallel inventory of supply and utilization of hospitals and other health service facilities.
- . An appraisal of educational resources available.
- A study of modifications required in patterns of medical education.
- An assessment of the economic resources available for health services and education.
- . Establishment of goals for health achievement and determination of manpower requirements for those goals.

Subsequent inquiry resulted in the selection of Colombia as the pilot country and, to provide the necessary "measurement of the health of the people and health services received", the identification of one of the studies as a national health survey, with both household and clinical examination components.

Thus while the U. S. Health Survey was developed as a more or less freestanding continuing statistical program to meet the needs of a variety of users the Colombian Health Survey was developed as a part of, or an instrument of, a one-time undertaking to establish national health goals, with particular orientation to the planning interests and responsibilities of the Ministry of Health and the schools of medicine.

An original purpose of the United States Health Examination Survey was to identify nonmanifest disease in adults, and thus to complement household survey data. In contrast, plans for the Colombian survey developed on the assumption that in a country with a low level of literacy, and a low level of health services, household interview surveys alone could not provide adequate information on health status. Thus the examination was planned as a rather ambitious health appraisal for all age groups, to include diagnoses based both on clinical and laboratory findings, with some emphasis on acute and parasitic conditions, and on malnutrition.

SURVEY DESIGN

In the United States, the designs of the independent Health Interviews Survey and Health Examination Survey are based on the principle of optimum allocation of resources. Due to differing cost and variance configurations, the designs of the two surveys are quite different. The Health Interview Survey with low unit costs has a large sample of households (about 40,000 annually) in a large sample of places (over 300) throughout the United States. On the other hand, for the Health Examination Survey, for which unit costs were high, a relatively small number of adults were examined, namely 6,500, in a small sample of places (42).

In Colombia, the basic decision to collect information from health interviews and examinations to provide a composite picture of an individual's health status, both as he knew it and as it was known through a clinical examination, implied a small sample of places and persons. Since the unit cost for each person in a health interview survey is low relative to the cost for both interview and clinical examination data (in Colombia, \$1.40 as compared to \$26.20), it was possible to provide greater reliability for demographic, disability and medical resource utilization data at little extra cost by collecting interview data only for a larger sample of persons than would be both interviewed and examined in each place. The design therefore called for the interviewing of the persons in a sample of about 240 households in each of 40 sample places, either a city, part of a large city or a municipio. A subsample of one out of every 10 persons interviewed was to be examined. Of the 52,480 persons eventually interviewed in 8,960 sample households, 5,260 were selected to receive a clinical examination.

We cannot describe the Colombian sample design in any detail in this paper. A very complete description of the sample has been published in Spanish by the Colombian Health Survey. The design is not unlike that of the Health Examination Survey; it is a complex multistage probability design. Colombia, excluding the territories, vast undeveloped areas comprising 47.3 percent of total land area in Colombia but only 1.3 percent of the population, was first geographically divided into 716 primary sampling units, the units typically being a single municipio, equivalent in many ways to our counties. Because of their large population size, 10 of the 716 PSU's were selected in the sample with certainty. The remaining PSU's were classified into 30 strata according to their elevation, population density, and population size. Next, a sample of PSU's or places were selected from each strata using a controlled technique, the control criteria being geographic proximity and a 4-scale index of health services available. The controlled selection technique differed from other applications known to us in that it controlled for 2 half-samples as well as the total sample. The two half-samples later formed the basis for estimating the reliability of estimates by a pseudo-replication method. One half-sample of 20 sample places, a probability sample of Colombia, was scheduled early in the survey period so that national estimates could be provided if events in 1966 would not permit the completion of all 40 sample places. The probabilities of selection of PSU's were based on current population projections from the 1951 Census. At the risk of some bias, a ratio estimation to 1965 urban and rural

population projections from the 1964 Census compensated for deficiencies in the probabilities used in selecting primary sampling units from strata.

By the time we were ready to select the samples of clustered households, advance counts of the number of census schedules or buildings in each 1964 Census Enumeration District had been provided by the National Statistics Department for the 40 sample places. After a conversion of the listed information from buildings to housing units and segments of about 10 housing units, a random systematic selection technique drew a sample of 24 area segments, some urban and some rural, for each sample place. The design called for area clusters containing about 10 households because enumerator assignments in urban areas in the one-day "de facto" 1.64 Census averaged 10 households.

Persons to be examined, about 130 in each sample place, were selected from among the persons in sample households, excluding infants. To avoid the necessity of returning to sample households, a method was developed to enable the interviewer to select a sample of persons, stratified by age and sex, at the completion of each household interview. A structured field quality control program, including interviews, gives us a high degree of confidence that the field selection procedures and interviewing were carried out according to instructions.

Once the decision was reached to include clinical examinations in the Colombian National Health Survey, it then remained to devise a statistically sound and operationally feasible survey plan, and then carry it out. An approach discussed but rejected would have the seven medical schools in Colombia interview and examine representative samples drawn from the communities surrounding the schools. Beside the obvious drawback of improper representation of the Colombian population using such an approach, the examination findings would have been subject to large measurement errors arising from the use by the medical schools of varying examination methods, equipment, standards of procedure, and laboratory techniques. The directors of the Colombian survey were quick to realize the need for an operational plan which would reduce measurement errors in examination findings to a minimum. In general, the problem of measurement error can be attacked in two different ways. One approach is to increase the number of measurements, measurers, and measurement devices (the possible sources of measurement error) and by so doing obtain an expected value with a negligible variance. The other approach is to standardize the measurers and measurement devices. In interview surveys both approaches are generally used. Therefore, a standardized questionnaire is used by a number of interviewers trained to conform more or less to standard interviewing procedures.

The U. S. Health Examination Survey attempts to eliminate the measuring device as a source of measurement error by doing the examinations in a mobile health examination center (two were used by Cycle I) consisting of 2 or 3 specially built tractor-drawn trailers, containing X-ray, other examining equipment and laboratory facilities. The mobile center and its calibrated equipment consequently set the standard and provide little or no opportunity for deviation from the standard in any of the examinations made in the mobile center. The cost for this attention to measurement error is obviously high for it sharply reduces the number of places from which persons can be selected for examination and the total number of persons that can be included in an examination survey within a reasonable time period.

In Colombia a mobile examination center was also necessary, since a large portion of the population is remote from any health examination facility. To transport persons selected for the examination survey to the nearest adequate examining facility frequently would require overnight stays and in some instances air transportation. Such inconveniences probably would have resulted in high nonparticipation rates. However, the standardized mobile center approach used in the United States could not be followed in so mountainous a country with such an incomplete road system. Instead, a "standardized examining center" was constructed in each of the 40 sample places. Existing health centers, clinics, or hospitals in or near the sample places provided the walls. Most of the examining equipment and the examination technicians were moved from place to place by whatever transportation means was required. In some cases this meant a combination of air, land and water transportation. Five clinical teams and sets of standardized examining equipment were in operation throughout most of the 8-month survey period. In amazing feats of logistics each clinical team packed-up from one examining location, set up in the next location about μ days later, and after typically two weeks of examinations packed-up again.

DATA COLLECTION

Interviewers in the United States National Health Survey, all women, are part-time employees of the Bureau of the Census. No prior education or work-experience in the health fields is required of the interviewers. Since the housewife is usually the respondent to a household interview in the United States, it is believed that female interviewers obtain a higher degree of cooperation and more complete reporting. Contrastingly, interviewers in the Colombian Health Survey were medical students, most of them male. The traditional labor pool in the United States--educated housewives seeking part-time employment--does not exist in Colombia. Medical students were used in Colombia not because of any conviction that they would be superior interviewers, but because they represented the best available manpower supply. Since medical students could be away from classes for only 15 days at most, a new set of 6 interviewers were trained and employed in each of the 40 sample places. Another important, nonstatistical reason for

using medical students is closely tied to the origin and objectives of the Colombian Health Manpower and Medical Education Study. It was believed that participation of medical students as interviewers, and residents in internal medicine and pediatrics as clinical examiners. would offer a unique opportunity for medical students and young physicians to actually get out into the country and to see what people were like in their homes and what their health problems were in relation to the places in which they lived. The Colombian medical schools have since indicated that the kinds of perception gained by medical students in these field operations are already having substantial impact on the teaching in their institutions. As yet there is not sufficient evidence to evaluate the performance of medical students as interviewers. The high response rate (97 percent) indicates that there was little difficulty in gaining cooperation of sample households.

Turning to the questionnaire of the Colombian Survey, again we find that conditions peculiar to Colombia and perhaps other developing countries were as important in determining its composition as were the requirements for information. At the outset George Kearns, the consultant on interview survey methods, recommended against any wholesale adaptations of the Health Interview Survey questionnaire. The latter is an unusually detailed and complex questionnaire that requires weeks of instruction and practice to bring new interviewers up to the level of satisfactory performance. In Colombia both the interviewers and their supervisors were totally without prior interviewing experience. It was extremely doubtful that their level of performance could be brought up to that of their U.S. counterparts in a short time. So, while the Colombians retained many of the basic features of the Health Interview Survey questionnaire, they discarded one of its principal and most difficult elements -that series of questions which seeks to obtain detailed characteristics of each reported morbidity condition, subsequently coded to the 4-digit ICDA code. Instead, a modest effort was made to measure prevalence of only nine of the more common chronic conditions recognizable even by an untutored or illiterate layman.

The sections of the United States questionnaire retained with only minor modifications were those that obtained information on demographic and socio-economic characteristics; incidence of acute illness and injuries, with the standard two-week recall; restricted activity due to illness; "doctor" visits (and within that, visits to, or service by "quasi" health practitioners such as midwives, pharmacists, teguas, and so forth); amount and source of payment of the "doctor" bill; and dental visits. There was also a section on hospitalization which included questions on cause, length of stay, and amount and source of payment. The recall period for hospitalization was not the traditional "last 12 months"--a difficult concept at best, even here in the United States. Instead it began with a fixed almost universally celebrated and remembered date in Latin America, depending on

the month of interview either New Year's Day or Holy (Easter) Week. Because of the varying interview months for the 40 sample places, an adjustment factor was used to convert the data from the recall periods of from 9 to 15 months into annual estimates. As far as we know this caused little trouble in execution and hopefully reduced the response error experienced when using the "last 12 months" recall period.

As to the mechanics of the form, we like to think of it as a model of simplicity. We know from reports of the observers that it presented few difficulties even to the marginal interviewers. It was highly structured, easy to read, with uncomplicated skip patterns, and with sufficient space for recording answers. The format of the main part of the questionnaire was particularly efficient. It was a booklet with provision for a seven person family -- with facing pages for each person. Multiple events such as hospitalizations were handled by using additional "person" pages for each event. The left third of the page contained instructions to the interviewer, the central third contained the actual questions, while the right third was the answer section including a shaded code area for those categories which could not be pre-coded. This eliminated the necessity for separate transcription and simplified the punching operation. Undoubtedly some improvement could have been made in the questionnaire. However, despite the relative inexperience of the project staff, it was one that was carefully phrased, logically put together, and one which worked well both in the field and at the processing stage.

The clinical examination followed the U. S. protocols, except for the emphasis on acute and parasitic diseases and nutritional status. Procedures were carefully standardized, as in the United States survey. Special techniques were developed to assure accurate laboratory work. Because of the short supply of technicians, the Colombian National Institute of Health provided a 3-month training program for technicians for the clinical teams, and made the laboratory determinations for a number of the tests.

DATA PROCESSING

Up to this point our comparisons have probably been prejudiced towards those areas which cast the Colombian survey in a favorable light. As should be expected for an undertaking of such dimensions, some problems were encountered for which no satisfactory solution was found, and some mistakes were probably made in planning and executing the survey. Naturally, we tend to forget most of these. However, what was probably our greatest failure as senior consultants continues to haunt us today. To date only a fraction of the information contained in the Colombian National Health Survey has been tabulated, particularly with regard to the examination findings. Although two excellent reports on findings from the interview survey have been published, as have several methodological reports, progress has been slow. Our failure was one of not giving more emphasis to data processing requirements at an early stage, although we included a data processing consultant on the staff of U. S. consultants from the beginning.

The primary problem encountered in data processing was not one of sheer volume of data, but rather a lack of the advanced programming skills and tabulation control procedures necessary for wringing out correct tabulations from the raw data of a complex sample survey. This is not a problem peculiar to Colombia, but one which is widespread among developing nations, not to mention new programs in the United States. The U. S. National Center for Health Statistics has developed the needed programming skills and control procedures over a period of years of training, practical experience and a close working relationship with the Bureau of the Census. We allowed ourselves to believe that these skills had already been developed in the National Statistics Department (DANE), the initial data processing agent for the Colombian National Health Survey. We believe it is correct to say that as yet neither DANE nor the present contractor have developed to that stage. In the meantime programs of the U.S. National Health Survey have been modified and used in Colombia to produce marginal tables.

We feel that the problems we are now encountering could have been forestalled if: a person with basic programming skills had been added as a full-time member of the Colombian staff and trained for 3-months at the U. S. National Center for Health Statistics; and the entire Colombian staff had been better indoctrinated in the need for and methods of tabulation control procedures.

SUMMARY

At the October 1,67 Round Table meeting of the Milbank Foundation, convened to review the methodology of the Colombian National Health Survey and its implications for the social sciences and health planning in Colombia, one of the participants asked "Scientifically, what is new in this report?" He was referring to the findings, in his belief that they simply confirmed findings already well known. However, the question could also be directed at the methodology, and in a sense was because the methodology generated the findings. The answer to the question "what is new?" would then be the demonstration that national sample surveys can be carried out in a Latin American country despite very formidable obstacles. Among the many deep satisfactions we have gained from our Colombian venture is the knowledge that science has been advanced in Colombia and to some extent outside Colombia. The Colombians have been exposed to probability surveys, the methodology is now well-documented in Spanish for others to profit from, and a nucleus of experienced survey technicians has been formed. The methodology of the Colombian survey has been extended to a dental survey, now in the field in Venezuela, and has influenced the planning of health surveys in

Argentina and Brazil. Finally it has been shown that a program developed for national health planning purposes can also provide the basis for international comparisons in morbidity and health services.

REFERENCES

- Colombian National Health Survey: Planning, Methods, and Operation. Reprinted from the Milbank Memorial Fund Quarterly, April 1968, Vol XLVI, No. 2, Part 2, New York, N.Y.
- (2) Gomez, Luis C. and Losee, Garrie J.: La Muestra de la Investigacion Nacional de Morbilidad. Bogota, Minasalud, 1968.
- (3) Health Manpower and Medical Education in Latin America. Report of a Round Table

Conference. Reprinted from the Milbank Memorial Fund Quarterly, January, 1964, Vol XLII, No. 1, New York, N.Y.

- U. S. National Health Survey. Origin and Programs of the U. S. National Health Survey. Public Health Service Publication No. 584-Al, Washington, D C., 1958.
- U. S. National Health Survey. Plan and Initial Program of the Health Examination Survey. Public Health Service Publication No. 584-A4, Washington, D. C., 1962.
- (6) U. S. National Health Survey. The Statistical Design of the Health Household-Interview Survey. Public Health Service Publication No. 584-A2, Washington, D. C., 1958.

VIII

VOCATIONAL EDUCATION STATISTICS

Chairman, MORRIS B. ULLMAN, U. S. Office of Education

	Page
The Needs for Vocational Education Data - MORRIS ULLMAN, BERNARD MICHAEL and MARC MATLAND, U. S. Office of Education	137
Theoretical and Empirical Problems in the Analysis of the Economic Costs of Vocational Education - ERNST W. STROMSDORFER, TEH-WEI HU and MAW LIN LEE, Pennsylvania State University	144
Discussion - MORGAN V. LEWIS, Pennsylvania State University	153