SOCIAL AND ECONOMIC MORTALITY DIFFERENTIALS IN THE UNITED STATES, 1960: OUTLINE OF A RESEARCH PROJECT

Philip M. Hauser and Evelyn M. Kitagawa University of Chicago

The declines in death rates and increases in longevity have, without question, been among the most significant and dramatic consequences of man's activity as a culture-building animal. A high death rate is in the contemporary world one of the best indicators of a low level of economic development. In the economically advanced areas, death rates have been pared by approximately three-fourths in the modern era; and despite their older populations, economically advanced nations tend to have general death rates approximately one-fourth the level of death rates in the economically underdeveloped areas of the world.

The great declines in mortality have been the product of a combination of forces of increased productivity, political order, environmental sanitation, personal hygiene and modern medicine. Yet, despite the great advances that have been achieved, large mortality differentials are still to be found even in the most advanced nations. To obtain a comprehensive picture of these differences among the various areas and subgroupings of the population in the United States in relation to personal, social, economic, and housing characteristics is the objective of the research now under way which is described in this paper.

The findings of this study, in revealing the populations within the United States which have achieved the lowest and highest death rates, respectively, will point to the areas and groups in which substantial reductions in mortality can yet be achieved and, also, will indicate the levels to which death rates can be further reduced. The findings of this study thus can serve as guides to public and private health programs in continuing efforts to increase longevity.

Many studies have been conducted on the death rates of the population of the United States and of other nations.¹ But nationwide information on mortality differentials by social, economic and housing characteristics has been restricted by reason of the limited information provided by the official record of mortality. That is, the death certificate as the official local legal record of death, a copy of which is made available to the National Office of Vital Statistics for statistical purposes, is limited to a few descriptive items about the person and the cause of death. The only item on the death certificate which may be considered a measure of the social-economic status of the deceased person is occupation. But it has not been possible to do much in the analysis of mortality by occupation because of the great discrepancy between occupational return on the death certificate, which provides information

¹/Dublin, L.I., A.J. Lotka and M. Spiegelman, Length of Life (New York: Ronald Press, 1949). Spiegelman, Mortimer, <u>Significant Mortality and</u> <u>Morbidity Trends in the United States Since 1900</u> (Revised) (Philadelphia: American College of Life Underwriters, 1960).

for the numerator of the death rate, and that reported in the population census from which the denominator of the death rate is drawn. In fact, even in respect to the other items on the death certificate such as marital status and color or race, the death rate is subject to error because of the lack of correspondence in the information reported on the death certificate and census schedule, respectively. Because of these discrepancies in the reporting of occupation on death and census records, the preliminary findings of a 1950 study of mortality of white males 20-64 years of age are restricted to only five occupational levels. Even with this limitation, however, there is an inverse relation between occupational level and mortality incidence, the death rate of laborers being twice as high as that of professional workers.²/ A good measure of the error in 1950 occupational death rates calculated by this method will soon be available.<u>3</u>/

Important information has been obtained on mortality differentials by social-economic status through an indirect form of analysis possible in cities and metropolitan areas where census tract data are available. In census tracted areas, it is possible to measure differentials in mortality, utilizing the social and economic characteristics of the census tract in which the decedents resided. A study now being completed in Chicago, for example, estimates 1950 mortality rates by social-economic status for the City of Chicago by allocating 1950 deaths to the 935 census tracts into which the city is divided, and using the regular census tabulations of population by census tracts as the base.4/. In this study, ageadjusted death rates are calculated for several "income levels," and also for several "quality of housing" levels by assigning death and population data for each tract as a unit to one of the levels on the basis of the median family income, or the proportion of substandard dwellings, in the tract. This method although feasible for census tracted areas, cannot be used to analyze differential mortality for non-tracted areas and, therefore, the entire country. Moreover, the method is subject to the error of ecological

2/Moriyama, I.M. and L. Guralnick, "Occupational and Social Class Differences in Mortality," in Milbank Memorial Fund, <u>Trends and Differen-</u> tials in Mortality (New York, 1956).

3/The Scripps Foundation for Research in Population Problems has undertaken a study of the correspondence of reporting of occupation on death certificates and census schedules for a sample of persons who died in the four-month period following the 1950 census.

^{4/}Hauser, P.M. and E.M. Kitagawa, "Differential Mortality in Relation to Quality of Housing" (Paper prepared for Western Branch, American Public Health Association Program, June 2-5, 1959, San Francisco). correlation, in necessarily involving the assumption that a summary index of social-economic status for an entire census tract applies to individual residents of the tract.

The study which is described below is designed to provide information on differential mortality despite the difficulties in doing so arising from the way in which official death data and population census materials are made available.

The Design of the Study

The data for the study will be information on death certificates for approximately 400,000 deaths that occurred from May through August, 1960; and the returns for these decedents and the entire population as obtained in the 18th Decennial Census of the United States as of April 1, 1960, in the Censuses of Population and Housing. The total number of deaths which occurred in the nation during the four-month period was about 500,000. All deaths of white persons under 65 years of age, and deaths of all nonwhite persons will be included in the study. A random selection of half the deaths for white persons 65 years of age and older will be excluded from the study, since such a large proportion of all decedents are 65 or older. This will provide, in all, a universe of approximately 400,000 deaths, for each of which there will be available a copy of the death certificate.

The serious problem posed by the limitation of personal, social and economic information about the decedent on the death certificate, and the lack of correspondence between the reporting of particular characteristics on the death certificate and on the population schedule for the same person, will be resolved by matching each of the 400,000 death certificates to the 1960 census schedules. Such a matching of the death certificates to the census returns, moreover, will provide not only all of the census information for the decedent but, also, for the decedent's family, household and dwelling unit. The matching procedure assures, automatically, comparable information for the numerator and the denominator of the death rate.

The matching operation will involve the following steps:

<u>Step 1</u>. Xerox prints of the 400,000 deaths included in the study will be made from the microfilm copies of death certificates forwarded by the states to the National Office of Vital Statistics. These Xerox prints will be sent to the Census Bureau. A duplicate set of the National Office of Vital Statistics punched mortality cards for these deaths will also be forwarded to the Census Bureau.

<u>Step 2</u>. The Census Bureau will determine the 1960 E.D. (enumeration district) number of the place of residence of each deceased person, using the address shown in the Xerox print of his death certificate.

<u>Step 3</u>. The Census Bureau will search the 1960 census schedules for the schedule containing the name of each deceased person. The E.D. number assigned in Step 2 will direct the searcher to a file containing all of the census schedules for that E.D.

Step 4. The Census Bureau will collate the

information obtained from the death certificates and the census schedules, and will summarize this information on computer tape, which will be used to compile tabulations of matched deaths by characteristics of the deceased persons as reported on their April 1960 census schedules.

The major limitation in this procedure arises from the incompleteness of the matching which can be achieved. Past experience indicates that about 20 percent of the deaths will not be matched by reason of incomplete addresses on the death certificate, the mobility of the population resulting in the changing of address between the census date and date of death, incompleteness of census enumeration, illegibility of returns on death certificates and census schedules, etc. A matching of only 80 percent of the death certificates to the census schedules would undoubtedly produce biased results in understating death rates and in making questionable the representativeness of the decedents for whom the census information is successfully obtained. Fortunately, it has been possible to devise procedures which will eliminate or bring into reasonable control this possible source of bias.

To control this source of bias a special sample survey is being conducted in advance of the large-scale study in order to provide a basis for estimating the characteristics of the decedents for whom no census match will be possible. A questionnaire is mailed to the informant who provided the information on the death certificate for a sample of approximately 10,000 deaths for the same four-month period. The questionnaire calls for most of the information which was obtained in the 1960 census. To assure as high a response rate as possible provision was made to mail the questionnaires as soon as possible after the death records are received in the National Office of Vital Statistics. Provision was also made for follow-up requests by certified mail to persons not replying to the first mailing. Previous experience of the National Office of Vital Statistics indicates that information can be collected for about 85 percent of the deaths in such a mail survey. To control the possible bias of non-response at this point, personal interviews will be conducted with all, or an adequate sample of, non-respondents to the mail survey. It is anticipated that a response rate of about 90 to 95 percent will be achieved.

It may be gathered that approximately 20 percent (2,000) of the questionnaires collected in the sample survey will provide information about deceased persons not matched to the census schedules. Which fifth of the 10,000 decedents in the sample survey will represent unmatched cases will, of course, not be known until after the large-scale matching operation is completed. But the availability of the sample survey results will make possible, eventually, the estimation of the characteristics of the decedents for whom death certificates could not be matched to census schedules. Moreover, the sample survey results will include about 8,000 cases for whom the matching will have been completed. This information can be used for quality check purposes in analyzing the correspondence of information obtained on the death certificate, the census schedule and the survey questionnaire.

Needless to say, the handling of the individual death certificates and the census records in the matching operation will be done entirely by the staffs of the National Office of Vital Statistics and the Bureau of the Census. The confidentiality of the returns, both on the death certificates and the census records, will be scrupulously maintained. The participating governmental agencies after performing these processing operations will make available to the study directors, for analysis, the results of tabulations of magnetic tape containing the collated data. $\frac{5}{}$ Moreover, the base population statistics needed to calculate death rates will be tabulated by the Bureau of the Census in regular and special tabulations.

The Specific Objectives of the Study

The analytical objectives of the study are limited by the sampling design introduced into the 1960 Censuses of Population and Housing. For the characteristics which were collected in the Censuses on a 100-percent basis, mortality differentials can be calculated utilizing the full universe of deaths. The census characteristics collected on a 100-percent basis were, however, restricted to age, sex, color or race, marital status, relation to head of household (for family and household composition), tenure, occupancy status, condition of structure, number of rooms, type of housing unit, water supply, toilet facilities, bathing facilities, kitchen cooking equipment. All other items in the Population Census were collected on a 25-percent sample basis, and items in the Housing Census on a 20-percent or 5-percent basis. Thus, most of the population and housing information will be available for 25 percent of the decedents included in the study. As a result, there will be a sample of approximately 100,000 deaths, including about 80,000 matched deaths for whom the detailed census information will be available. Similar information will be available -- from the Special Sample Survey questionnaires -- for about 2,000 of the 20,000 deaths not matched with census schedules but presumably included in the 25-percent census sample. Detailed cross tabulations of mortality by social and economic characteristics of the decedent will, therefore, be affected by sampling variance, and especially will this be true for cause of death tabulations. Despite this limitation, however, it is clear that for the more important specific causes much valuable information can be obtained.

With these restrictions in mind the anticipated specific objectives of the study can be outlined as follows:

1. Geographic differences in mortality.

^{5/}The study is being conducted through the Population Research and Training Center at the University of Chicago. Principal investigators are Philip M. Hauser and Evelyn M. Kitagawa. The study design is the product of consultation with a number of people, including Iwao Moriyama, Lillian Guralnick and Monroe Sirken of the National Office of Vital Statistics, Charles Nam of the Census Bureau, and Harold Dorn and Mortimer Spiegelman.

Mortality differentials will be obtained for meaningful geographical areas for which analysis is not now possible by reason of the difficulties of achieving comparability in the classification of geographic areas from residence as reported on the death certificate and census schedule, respectively. The use of the geographic classification coded in the 1960 Census records, will permit the calculation of death rates for urban and rural areas as defined in 1960, for Standard Metropolitan Areas subdivided into central cities and rings, for urbanized areas, for groups of cities classified by size of population, and for farm areas. Thus, mortality analysis will be available for the new and significant areal categories for which rich census materials are available.

2. Nativity and ethnic differentials in mortality. The place of birth of the person and of his parents will be used to calculate death rates for first and second generation groups, by ethnicity. It will, therefore, be possible to analyze in a comprehensive manner for the nation as a whole not only the extent to which ethnic groups differ in mortality but, also, in at least broad categories ethnic mortality by cause of death.

3. Income differentials in mortality. Differential death rates by income will be possible on the basis of the 1959 income of the individual decedent, and the family income of the family of which he was a member. For example, mortality rates by family income can be computed for all those living in family groups (92 percent of the population in 1950); mortality rates by individual income can be computed for unrelated individuals (those not living in family groups); mortality rates by individual income can be computed for all men and women with an income; mortality rates for wives (and children) can be computed by income of husband (father).

4. Occupational differentials in mortality. The occupation and "class of worker" for each person at work during the week preceding the 1960 Census will be available and, also, the last occupation held by persons not at work the week before the census but who have worked some time during the past ten years. This information can be used to obtain mortality rates for both men and women with work experience since 1950; and mortality rates for women (and children) by occupation of husband (father).

Differential death rates by occupation will, in the main, reflect social and economic status differentials of which the occupation is an index rather than "occupation mortality risk." Although it may not be feasible, given the number of deaths in the study, the possibility of obtaining some measure of mortality risks for certain kinds of occupations will be investigated. The list of such occupations will be investigated. The list of such occupations would, of course, be confined to those in which long apprenticeships, restrictions for admission, and long tenure permit the conclusion that it is one in which the person has spent all, or an appreciably large part, of his working career.

5. Education differentials in mortality. For persons 25 years of age and older, the census information on years of school completed can be used to calculate differential death rates by educational level. 6. Mortality differentials by family status. Census information will permit the calculation of death rates not only by marital status of the deceased but, also, by family status of the deceased, including presence and age of spouse, and number and age of surviving children; by parity (number of children ever born) for all women who have ever been married.

7. Mortality differentials by housing and environmental factors. Death rates can be computed by various characteristics of the dwelling unit of the decedent, including rent, value of dwelling unit and quality of housing. To some extent, thus, environmental factors can be related not only to mortality in general but, also, to cause of death.

8. Social and economic differentials in mortality for selected causes of death. To the extent that the size of the sample permits, deaths by each of the social and economic variables described above, will be studied separately for selected causes of death. This will include, for example, income differentials in cancer mortality, occupational differentials for deaths from diseases of the heart, etc. It is anticipated that some social-economic differentials in the causes of mortality can be obtained for at least the 10 to 15 leading causes of death.

9. Interrelationships among the social and economic factors. The interrelationships among the various social and economic factors will be explored by analyzing mortality differentials with respect to each factor holding constant one or more of the other factors. For example, occupational differentials in mortality will be analyzed holding income constant, and ethnic differentials in mortality will be analyzed holding income constant. Similarly, the effect of income will be analyzed while holding occupation and education constant, etc.

10. Methodological studies. The investigation will permit a number of methodological studies. One to which reference has already been made relates to variability in response, possible by comparing the returns for the same person on the census schedules, the death certificates, and the special sample survey. It will also be possible to study the ecological correlation problem mentioned earlier, by comparing social and economic differentials in mortality based on direct cross tabulations of a sample of decedents, with results obtained from the same sample using the social and economic characteristics of the census tracts in which they resided. At a number of points operational studies can be made on the linkage of records from independent sources, which can throw light on the feasibility of matching studies of various types of documents.

Aspects of the Analysis

In addition to comparisons of general mortality rates for the diverse geographical areas and population groupings possible, various types of refined mortality measures will be employed. For one thing, mortality analysis will be made, as far as possible, for significant age groupings, namely, for decedents 25 to 44 years, 45 to 64 years, and 65 years of age and older, respectively. Analysis will also be made of mortality at younger ages but, because of problems affecting the denominator for infants, there is not much that can be done in this study in the way of infant mortality analysis.

In addition to analysis for the significant age groupings, age-adjusted mortality measures will be used throughout the study. Furthermore, largely for purposes of graphic and convenient summarization, abridged life tables will be calculated by a number of characteristics not hitherto possible, for example, by education, income, ethnicity, marital status and the like.

To highlight differential death rates, mortality ratios will be calculated relating death rates of subgroupings to the general or lowest death rates used as a base.

Two types of adjustment will affect the results of the entire investigation. One will involve the adjustment of the data for the unmatched deaths. It will be possible to estimate the characteristics of the decedents for whom certificates have not been matched to the census schedules, using data from the Special Sample Survey for a sample of these decedents. With such an adjustment, mortality rates can be calculated so as to make the results representative of the entire United States.

Finally, since the deaths are for four months of the year, the problem of differential mortality by season must also be taken into consideration. For the universe as a whole, seasonal adjustment can readily be made by relating mortality for the four-month period of observation to the national mortality data available for the year as a whole. Seasonal adjustments by characteristics, however, will, of course, be restricted to the items that are available on the death certificate itself, including age, sex, color or race, marital status, cause of death, place of residence. Problems of seasonality may conceivably affect differential mortality by various other social and economic factors. Preliminary investigation, however, indicates that serious distortions as a result of this problem are unlikely.

Phasing and Cooperation

This study of differential mortality is necessarily a cooperative venture. It would not be possible without the unstinted cooperation and interest of the National Office of Vital Statistics, the state and local Vital Statistics Offices and the Bureau of the Census. These agencies have indicated their interest in participating in the investigation and have already, in the essential preliminary work, enthusiastically provided the necessary assistance.

The entire study was made possible by a grant from the U.S. Public Health Service and the National Institutes of Health.⁽²⁾ The grant is for a five-year period which began in May, 1960. It is expected that the coding, matching and collation of the data will be completed by the end of 1962, and tabulations should be available by the end of 1963. This leaves the

^{6/}The total amount of the grant is \$1,017,000, about 90 percent of which will be subcontracted to the National Office of Vital Statistics and the Bureau of the Census to collect the basic information.

120

principal investigators approximately two years for the analysis and writing of the reports in which the results will be published.

Needless to say, in a comprehensive and large scale study of this type, a number of supplemental and special investigations will be possible. Various agencies and individuals have already indicated an interest in special tabulations and analytical possibilities. The investigators themselves have in mind a number of extensions of analysis of the materials which may become desirable as the study reaches its later stages.

The investigators regard the data which will be collected as being largely in the public domain and are prepared to cooperate as best they can in helping to make the data available for special studies, within the limits of available resources and sharing of costs. Further reports on the status of this study will be made from time to time.