Journal of Official Statistics, Vol. 11. No. 3, 1995, pp. 233–260

Sources of Data on Socio-Economic Differential Mortality in the United States

Donna L. Hoyert, Gopal K. Singh, and Harry M. Rosenberg[1]

This paper reviews several data sets that are available to assess mortality patterns in the United States and to examine changes in the health and social status of identifiable segments of the U.S. population. The specific data sets described: (1) the National Vital Statistics Mortality Data File, (2) the National Mortality Followback Survey, (3) the NHANES Epidemiologic Followup Study, (4) the Social Security Administration/ Current Population Survey Matched Files, and (5) the U.S. National Longitudinal Mortality Study, reflect different data collection methodologies and can support different analytical techniques. Selected results from previously published and unpublished studies are presented for a few variables to illustrate the analytic potential of socio-demographic variables available for research on mortality differentials.

Key words: Vital statistics; mortality; data sources.

1. Introduction

European research has a strong tradition of examining socio-economic differentials of mortality (Kunst and Mackenbach 1994; Valkonen 1993; Valkonen, Javanainen, Koskinen, Martelin, and Martikainen 1991; Caselli 1991; Fox, Goldblatt, and Adelstein 1982; Desplanques 1976; Kristofersen 1979; Valkonen 1982; Eriksen 1983). Analytical work in the United States has focused on racial differences, but in recent years, includes behavioral and socio-economic factors in mortality research on a more regular basis (Rogers 1992; Pappas, Queen, Hadden, and Fisher 1993; Sterling, Rosenbaum, and Weinkam 1993; Sorlie, Backlund, and Keller 1995). This paper reviews national mortality data bases available as tools for assessing the health of Americans and for understanding shifts in the health and social status of identifiable segments of the U.S. population. Description of the data sources will indicate the data collection methodology, demographic and socio-economic content, analytical methods that can be used, and a few illustrative examples.

In the U.S., use of mortality data to set goals for and to monitor the health and social situation of the nation and geo-political areas such as states, counties, and communities is growing, in part, as a result of the Federal Government's initiatives in disease prevention and health promotion. In U.S. Department of Health and Human Services (1991), U.S. health goals and objectives for the year 2000 are listed in 22 broad categories (Figure 1). Within the broad categories are more than 300 hundred

[1] Centers for Disease Control and Prevention, National Center for Health Statistics, 6525 Belcrest Road, Room 840, Hyattsville, MD 20782, U.S.A.

1. Physical activity and fitness
   1.1. Reduce coronary heart disease deaths to no more than 100 per 100,000 people
2. Nutrition
3. Tobacco
   3.3. Slow the rise in deaths from chronic obstructive pulmonary disease to achieve a rate of no more than 25 per 100,000
   3.4. Reduce cigarette smoking to a prevalence of no more than 15% among people aged 20 and older
4. Alcohol and other drugs
   4.2. Reduce cirrhosis deaths to no more than 6 per 100,000 people
5. Family planning
6. Mental health and mental disorders
   6.1. Reduce suicides to no more than 10.5 per 100,000 people
7. Violent and abusive behavior
   7.11. Reduce by 20% the proportion of people who possess weapons that are inappropriately stored and therefore dangerously available
8. Educational and community based programs
   8.2. Increase the high school graduation rate to at least 90%, thereby reducing risks for multiple problem behaviors and poor mental and physical health
9. Unintentional injuries
   9.14. Extend to 50 States laws requiring safety belt and motorcycle helmet use for all ages
10. Occupational safety
    10.7. Reduce to no more than 15% the proportion of workers exposed to average daily noise levels that exceed 85 dBA
11. Environmental health
12. Food and drug safety
    12.2. Reduce outbreaks of infections due to Salmonella enteritidis to fewer than 25 outbreaks yearly
13. Oral health
    13.8. Increase to at least 50% the proportion of children who have received protective sealants on the occlusal (chewing) surfaces of permanent molar teeth
14. Maternal and infant health
    14.1. Reduce the infant mortality rate to no more than 7 per 100,000 live births
15. Heart disease and stroke
    15.2. Reduce stroke deaths to no more than 20 per 100,000 people
16. Cancer
    16.3. Reduce breast cancer deaths to no more than 20.6 per 100,000 women
17. Diabetes and chronic disabling conditions
    17.9. Reduce diabetes-related deaths to no more than 34 per 100,000
18. HIV
    18.4. Increase to at least 50% the proportion of sexually active, unmarried people who used a condom at last sexual intercourse
19. Sexually transmitted diseases
20. Immunization and infectious diseases
21. Clinical preventive services
    21.3. Increase to at least 95% the proportion of people who have a specific source of ongoing primary care for coordination of their preventive and episodic health care
22. Surveillance and data systems

Fig. 1. Year 2000 objective areas and selected examples of objectives

specific objectives to reduce preventable death and disability, improve quality of life, and reduce health differentials within the population (National Center for Health Statistics (NCHS) 1993a).

Mortality data from the vital statistics system are used to monitor progress in reaching many of the goals listed in U.S. Department of Health and Human Services

(1991) and are a widely used source of mortality information by the public, policy-makers, and researchers. In fact, mortality data from the vital statistics system have been a key source of health status indicators for the U.S. and other countries since the 1930s. However, additional U.S. national mortality data sources have been developed over the years. The data sources differ in the variables available for analysis, data collection methodologies, and support different analytical approaches.

1.1. Analytical methods

A variety of analytical methods are used in mortality analysis. The simplest of these methods are such descriptive measures as frequencies, proportions, ratios, and crude death rates. The crude death rate is defined as the number of deaths during a given year divided by the number of person-years during the same period, which is generally the midyear population. Because the risk of mortality varies considerably by age, a more refined measure of mortality is the age-specific death rate, which is defined as the number of deaths in a specific age group divided by the person-years of exposure in that age group. Age-specific death rates are, however, a cumbersome and inefficient method to analyze mortality trends and differentials if a number of age groups are involved. The standardized death rate, which controls for the effect of age composition, is a summary index of mortality that succinctly describes a given age-schedule of death rates. The directly standardized death rate, also referred to as the age-adjusted death rate, is the weighted average of the age-specific death rates, with weights being the age composition of a standard population.

In a longitudinal design in which a population exposed to the risk of death is followed through time, thereby allowing the determination of the exact length of follow-up or the time elapsed since death, an even more refined measure – instantaneous death or hazard rate – can be used to analyze mortality differentials. These hazard rates can be modeled in a survival regression framework as a function of time varying or time independent covariates. Commonly used survival regression models in mortality analysis include Cox proportional hazards model, discrete-time logistic model, and Weibull, Gompertz, and Poisson models.

1.2. Different approaches to collect data

The data sources described in the paper reflect a range of approaches available to collect data including use of existing administrative records; personal, mail, and telephone interviews; and clinical examinations. In some cases, several approaches are used in the same study.

Administrative records have the advantage of routine standardized collection and can be cost-effective and require little time per record to collect. Unit non-response is typically low because of legal requirements for maintaining records on various events. For example, a death certificate must be filed before a body can be buried or cremated (NCHS 1989). This type of data collection is often best for learning about rare events or events in small geographic areas because no sampling is done; instead, records are maintained for all events.

Survey methods are generally used to collect more extensive batteries of questions

Source: National Center for Health Statistics, U.S. Government

Observation plan: Period data

Data collection instrument: State death certificates

Period of data availability: 1900–present

History of data collection: The first death statistics published by the Federal Government concerned events in 1850 and were based on statistics collected during the 1850 Decennial Census. In 1880 a national "registration area" was created for recording deaths reported on death certificates. In 1900 the system was formalized in anticipation of the Census Office becoming a permanent entity, routine publication began

Coverage: In 1880, the registration area consisted of two states, D.C., and several large cities having efficient systems for death registration, the death-registration area continued to expand until 1933, when, for the first time, it included the entire United States. It is believed that more than 99% of deaths occurring in the U.S. are registered

Variables: Geographic, demographic, social, and medical conditions and diseases contributing to death (see Table 1)

Number of records: 2,169,518 deaths in 1991

Methodology supported: Frequency; proportion; ratio; crude death rate; age-specific death rates; standardized death rates

Shortcomings: Limited set of variables; another data source must be used to calculate the population at risk

Selected references: Rosenberg, Burnett, Maurer, and Spirtas 1993; NCHS 1993b

Short abstract of main findings: Example: Rosenberg et al. (1993) presented data on mortality by occupation and industry for 12 states in 1984. They identified links between specific causes of death and occupations and industries for males and females. As an illustration, males in mining occupations had elevated mortality risks relative to the expected risk for three causes of death: chronic obstructive pulmonary diseases and allied conditions, pneumoconioses and pnemopathy due to inhalation of other dust, and accidents mainly of industrial type

Fig. 2. National Vital Statistics Mortality Data File (VSUS)

on specific topics for a sample of the population. More effort is expended on collecting quality data through survey design and use of trained interviewers but surveys are more costly and time-consuming to complete. Unit non-response can be a problem with surveys although government sponsored surveys have less problem with unit non-response. The major vehicles through which surveys are administered are the personal interview, mail questionnaire, and telephone interview. There are trade-offs between cost, response rate, reaching geographically dispersed populations, collecting detailed information, and applicability for subgroups of the population associated with each vehicle (Babbie 1986; Groves 1989).

Clinical examinations offer an opportunity to gather information that a subject might not know about or might not be routinely collected in administrative records such as a medical record. This type of data collection can be standardized but requires medical equipment and a trained staff to collect, must be accessible to study subjects, is costly, and has problems with unit non-response.

Analytical potential of data sources can be enhanced by linking and merging data collected but this raises a number of legal and ethical questions, foremost of which,

Source: Directed by the National Center for Health Statistics, U.S. Government: jointly undertaken with a number of agencies of the U.S. Government

Observation plan: Period data

Data collection instrument: Mail questionnaires

Period of data availability: 1960, 1961, 1962–1963, 1966–1968, 1986

History of data collection: The first mortality followback survey was conducted by the National Center for Health Statistics in 1960. Periodic surveys have been conducted since then; the latest survey is currently in the field. Co-sponsoring agencies differ in each specific survey

Coverage: Systematic sample of deaths in the United States. Some deaths were selected with certainty by cause, race, or ethnicity. The overall response rate for the survey in 1986 was 89%

Variables: Demographic, social, economic, medical care utilization, health in the last year of life, behavioral, and medical conditions and diseases contributing to death (see Table 1)

Number of records: 16,589 completed questionnaires in 1986

Methodology supported: Frequency; proportion; ratio; crude death rate; age-specific death rates; standardized death rates

Shortcomings: Limited geographic detail; another data source may be needed to calculate relative risk

Selected references: Pappas et al. 1993; Powell-Griner and Rosenberg 1991

Short abstract of main findings: Example: Pappas et al. (1993) found age-adjusted death rates decrease with increasing levels of education and income

Fig. 3. National Mortality Followback Survey (NMFS)

concerns confidentiality (Ashley, Cole, and Kilbane 1991; Brackstone 1989; Kovar 1989). Data sources are merged but a variety of legal agreements are involved depending on the particular data sources in question. The data sets described in this paper do link together different records but identifying information that would enable a data user to merge other records are stripped from the data files before release. Many of the data sources described in this paper use the National Death Index program. This is a popular service provided by the Federal Government to assist in tracking study participants who have subsequently died. The service provides the following information to researchers: the state where the death occurred, date of death, and death certificate number. This information is only provided to approved research projects in which the researcher was able to provide enough identifying information to compare with National Death Index information. The researcher must then approach each of the state vital statistics offices about purchasing copies of the death certificates or statistical data from the death certificates depending on state laws on public availability of death certificates (NCHS 1990).

2. Data Sources

2.1. Overview of mortality data sources and allied data sources

Each of the data sources described in this paper, (1) the National Vital Statistics Mortality Data File (VSUS), (2) the National Mortality Followback Survey (NMFS),

Source: Directed by the National Center for Health Statistics (NCHS), U.S. Government: jointly undertaken by NCHS and the National Institute on Aging in collaboration with other agencies of the Public Health Service, U.S. Government

Observation plan: Longitudinal data

Data collection instrument: links personal interview survey; clinical examination; death certificates; nursing home records; telephone survey using CATI

Period of data availability: 1971–1975; 1982–1984; 1986; 1987

History of data collection: The study was designed to investigate relationships between clinical, nutritional, and behavioral factors assessed in the first National Health and Nutrition Examination Survey; subsequent morbidity, mortality, and hospital utilization; changes in risk factors, functional limitation, and institutionalization. The last follow-up was conducted in 1992 but data have not been released

Coverage: All subjects, aged 25–74 years, receiving a clinical examination in NHANES I, a nationally representative sample of the civilian non-institutionalized population, were selected for this longitudinal study; in the 1982–84 follow-up, 93% of eligible subjects were traced. Interviews were obtained for 93% of survivors and proxy interviews were completed for 84% of those deceased in 1987

Variables: Clinical examination, nutritional, demographic, social, behavioral, health care utilization (see Table 1)

Number of records: 14,407 subjects followed through 1987; 3,212 were deceased by 1992

Methodology supported: Frequency; proportion; ratio; crude death rate; age-specific death rates; standardized death rates; instantaneous death or hazard rate with covariates including time-varying covariates

Shortcomings: Small number of deceased sample members at present; loss of eligible subjects in reinterviews; institutionalized population not included in initial sample, limited geographic detail

Selected references: Feldman et al. 1989; Makuc et al. 1990

Short abstract of main findings: Feldman et al. (1989) found tangible socio-economic differentials in mortality for older adults during the period 1971–1984. In the case of heart disease deaths, socio-economic differentials were found even when controlling for smoking, blood pressure, cholesterol, and weight

Fig. 4. NHANES Epidemiologic Followup Study (NHEFS)

(3) the NHANES Epidemiologic Followup Study (NHEFS), (4) Social Security Administration/Current Population Survey Matched Files (SSA/CPS), and (5) the U.S. National Longitudinal Mortality Study (NLMS), is useful for investigating socio-economic differentials in mortality. Figures 2 through 6 summarize features of the data sets described in this paper.

The National Vital Statistics Mortality Data File (VSUS) is the data source used for most national mortality statistics published by the U.S. Government, and for international comparisons of mortality statistics. The vital statistics mortality data file compiles information reported on death certificates of every death occurring in the United States each year. Its primary advantages are its size, representation of decedents, geographic detail, availability back to the beginning of the century, and the frequency with which it is compiled. Its chief disadvantages are a limited set of variables, another data source must be used to calculate the population at risk, and

Source: Social Security Administration, U.S. Government

Observation plan: Longitudinal data

Data collection instrument: Link 1973 Current Population Survey; 1973 Internal Revenue tax records and Social Security Administration records; 1973–1978 Social Security earnings and beneficiary records

Period of data available: 1973–1978

History of data collection: In the 1970's, the Social Security Administration expanded the information in their own administrative data files by merging social security records with data obtained in the current population survey. Access to this and other Social Security Administration linked data files depends on tax regulations in the U.S.

Coverage: The study population represents a probability sample of households from the civilian non-institutionalized population of the U.S. The response rate to the CPS is usually around 96%. The coverage of Social Security death reporting increased over time. For deaths in 1977, 98.5% were recorded in the Social Security records

Variables: Demographic, social, economic (see Table 1)

Number of records: 13,900 reported in published studies (Duleep 1989)

Methodology supported: Frequency; proportion; ratio; crude death rate; age-specific death rates; standardized death rates; age-specific death rates; instantaneous death or hazard rates with covariates including time-varying

Shortcomings: Underreporting for persons who were not covered by Social Security (i.e., women, Black persons, and some white males under age 65 years)

Selected references: Duleep 1986; Duleep 1989

Short abstract of main findings: Example: Duleep (1989) found that the relative mortality levels were higher for those with less education and for those with lower incomes

Fig. 5. Social Security Administration/Current Population Survey Matched Files (SSA/CPS)

it cannot support more sophisticated analytical techniques such as time-varying covariates.

In the United States, registration of deaths is governed by state rather than Federal law. However, the Federal Government is required to collect and publish national vital statistics data. The data are obtained from the state system through a contractual arrangement between the states and the Federal Government under which the Federal Government shares the cost of collecting and processing vital records, including quality control activities. The Federal Government provides detailed specifications for producing the data, including coding and computer editing, to ensure uniformity and to promote the quality of the data produced by the states. The Federal Government also develops materials to assist states in training data providers (funeral directors, physicians, medical examiners, and coroners), but the principal responsibility for data collection, data processing, data quality maintenance and improvement, rests with the states (NCHS 1989).

The data are collected on a continuous basis with an administrative record, the state death certificate. To promote uniformity, the Federal Government recommends the *U.S. Standard Certificate of Death* as the model for the states to use in developing state death certificates (NCHS 1991). The medical certification of death section of the

Source: The National Heart, Lung, and Blood Institute; U.S. Government with the help of the Bureau of the Census and the National Center for Health Statistics, also U.S. Government

Observation plan: Longitudinal data

Data collection instrument: Links Current Population Survey telephone and personal questionnaire; state death certificate

Period of data availability: 1979–1985

History of data collection: The National Heart, Lung, and Blood Institute designed this study to provide a data source for examining socio-economic factors on mortality, particularly cardiovascular mortality. In the study, cross-sectional data are matched to data on deaths from 1979 forward. Data have been collected through 1989 but have not been released yet

Coverage: The study population represents a probability sample of households from the civilian non-institutional population of the U.S. The response rate to the CPS is usually around 96%. A two-step procedure was used to match the CPS records with death certificates to improve matching

Variables: Demographic, social, causes of death (see Table 1)

Number of records: 1,281,475 subjects being followed; in 1985, 44,828 were deceased

Methodology supported: Frequency; proportion; ratio; crude death rate; age-specific death rates; standardized death rates; instantaneous death or hazard rates with time-fixed covariates

Shortcomings: Lack of geographic detail; subjects initially are members of the civilian non-institutionalized population; variables limited

Selected references: Rogot et al. 1992a; Rogot et al. 1992b; Sorlie, Backlund, and Keller 1995

Short abstract of main findings: Example: Rogot et al. (1992a) found that mortality levels were elevated for persons who were not married, unemployed, had less education, or had lower income

Fig. 6. U.S. National Longitudinal Mortality Study (NLMS)

death certificate is standardized by an international agreement, embodied in the International Classification of Diseases (ICD) of the World Health Organization (World Health Organization 1977). Most state certificates conform closely to the standard, with modifications to meet particular state needs or legislation (NCHS 1991). There is a large number of studies on the validity of specific causes of death reported on death certificates which do indicate differential validity by cause but the body of research has not yielded a straightforward answer to the question about the quality of cause data (NCHS 1989; Gittelsohn and Royston 1982).

Data representing the population at risk for vital statistics events are obtained from the Decennial Census in the census years and intercensal estimates for years between the census. The U.S. Decennial Census collects data on social and economic factors of the population as well as enumerating the population as required in the U.S. Constitution. The census is conducted in years ending in "0," principally using a mail questionnaire, and is referenced to April 1 of that year. Two forms of the questionnaire are administered: a short form for most of the population and a long form with more detailed questions for a sample of the population (Robey 1989). The intercensal estimates for some characteristics such as educational attainment make use of more current data collected in the Current Population Survey (CPS). The CPS is a nationally representative household and telephone interview

survey of the civilian non-institutionalized population of the U.S. conducted monthly by the Bureau of the Census primarily to provide national estimates of employment (U.S. Bureau of the Census 1978); however, it contains a wide range of socio-economic and demographic characteristics.

The National Mortality Followback Survey (NMFS) is a data source that provides much greater detail on the characteristics and circumstances of decedents than is available from the death certificates which serve as its sampling frame. Its main advantages are the detail available on the last year of life, oversampling for certain causes of death and subpopulations, representation of the institutionalized population, and as an evaluation tool for death certificates. Its disadvantages are limited geographic detail, another data source is often used to calculate the population at risk, and it does not support all available analytical methods. The sampling frame for the NMFS is the Current Mortality Sample (CMS), a 10% systematic sample of death certificates received each month in the vital statistics offices in the states (NCHS 1993c). This cross-sectional survey is conducted periodically by the U.S. Federal Government by mailing questionnaires about six months after death to the next-of-kin or person listed on the death certificate as providing personal information for the decedent's death certificate (Seeman, Poe, and Powell-Griner 1993). For selected years, data from the death certificate and questionnaire are supplemented by data from medical facilities used by the decedent in the year prior to death.

Data representing the population at risk for the NMFS are commonly obtained from the National Health Interview Survey (NHIS). This is a continuing national survey of the U.S. civilian non-institutionalized population conducted in personal interviews. One member of each sampled household is interviewed to obtain information about the health and other characteristics of each living member of the sample household, primarily about health and morbidity (NCHS 1994; Kovar 1989). Another source of data to represent the population at risk could be used to address questions about the institutionalized population.

The NHANES Epidemiologic Followup Study (NHEFS) is a longitudinal data source for examining the relationship of baseline clinical, nutritional, and behavioral factors to subsequent morbidity and mortality. The advantages of this data source are that the population at risk is the initial sample population; clinical examination data are collected; data are collected in four reinterviews; and in most cases, the subjects respond for themselves in the baseline survey. All analytical methods are supported by this data. The disadvantages include a relatively small number of deceased sample members, loss of eligible subjects in reinterviews, no representation of institutionalized population in the initial sample, and lack of geographic detail. The data were collected by the Federal Government in the following stages: the first National Health and Nutrition Examination Survey (NHANES I) and personal interview data and administrative data collected in the subsequent follow-up in 1982–1984 and telephone survey data collected in 1986, 1987, and 1992 (Kovar 1989; Feldman, Makuc, Kleinman, and Cornoni-Huntley 1989; NCHS 1994).

The Social Security Administration/Current Population Survey Matched Files (SSA/CPS) is a longitudinal data source that can be used to monitor socio-economic differentials in mortality on a continual basis. This data set was created by linking

Table 1. Variables included in data sources

[Data Sources: National Vital Statistics Mortality Data File (VSUS), 1989; U.S. National Mortality Followback Survey (NMFS), 1986; NHANES Epidemiologic Followup Study (NHEFS), 1971–1984; U.S. National Longitudinal Mortality Study (NLMS), 1979–1985.]

Variable	Data sources					
	VSUS	NMFS	NHEFS	NLMS	Denominator data sources	
					Census/CPS	NHIS
Year of death	X	X	X			
Month of death	X	X	X			
Day of week	X	X	X	X		
State where death occurred	X	X	X			
County where death occurred	X		X			
Size of county of occurrence	X					
Region where death occurred	X					
Division where death occurred	X	X				
State of residence	X	X	X		X	
County of residence	X				X	
Size of county of residence	X				X	
City of residence	X				X	
Size of city of residence	X				X	
Metropolitan status of county of residence	X	X		X	X	X
Region of residence	X				X	X
Division of residence	X	X			X	
PMSA/SMSA of residence	X				X	
Sex	X	X	X	X	X	X
Race	X	X	X	X	X	X
Hispanic origin	X	X	X	X	X	X
Veteran status		X			X	X

Age	X	X	X	X	X	X
Marital status	X	X	X	X	X	X
State of birth	X	X		X	X	
Foreign born	X	X				X
Years lived in U.S.						X
Citizenship					X	
Native language					X	
Education	X	X	X	X	X	X
Income		X			X	X
Value of possessions		X				
Poverty status					X	
Insurance			X			
Kind of business or industry	X	X		X	X	X
Occupation	X	X	X	X	X	X
Work history					X	
Place of work					X	
Hours worked					X	
Living arrangements			X			
Household composition		X	X			X
Information on children in household		X	X			X
Number of children in family		X				
Children ever born			X		X	
Family size						X
Relationship of informant		X	X		X	X
Information on spouse		X	X			
Underlying cause of death	X	X	X	X		
Multiple causes of death	X	X				

Table 1. Continued.

Variable	Data sources					
	VSUS	NMFS	NHEFS	NLMS	Denominator data sources	
					Census/CPS	NHIS
Autopsy performed	X	X				
Place of accident	X	X				
Survival			X	X		
Transportation to work					X	
Travel to work					X	
Vehicle occupancy					X	
Mobility status					X	
Exercise			X			
Weight			X			
Fertility history			X			
Contraceptive history			X			
Functional status		X	X		X	
Alcohol and tobacco use		X	X			
Health in last year of life		X	X			
Diagnosis of specific chronic conditions		X	X			
Acute and chronic conditions			X		X	X
Health and utilization variables		X	X			X
Medical care in last year of life		X	X			
Medical care in recent period			X			X
Hospitalization in recent period			X			X
Hospital/nursing home stays		X	X			

data from two different Federal Government data sources: the 1973 Current Population Survey and the 1973 Internal Revenue and Social Security Administration records. Subsequent information was added to the baseline data for deaths occurring between 1973 and 1978 from Social Security earnings and beneficiary records (Duleep 1989). The main advantages of this data are the detailed information on income and employment, the population at risk is the initial sample population, and in most cases, the subjects respond for themselves in the Current Population Survey. Each of the analytical methods is possible with this data set. The major disadvantage of the data is that deaths are underreported for persons who were not covered by Social Security, in some way, prior to 1978. This was true, in many cases, of women, Blacks, and some white males under age 65.

The U.S. National Longitudinal Mortality Study (NLMS) is another longitudinal data source that maintains a continuous data base for examining socio-economic, occupational, and demographic factors associated with mortality in the United States by linking different Federal data records. The primary advantages of this data source are a relatively large size, the initial sample population makes up the population at risk, it can be used to validate death certificate data, and in most cases, the subjects respond for themselves in the base survey. All analytical techniques are applicable with this data source but the covariates are time-fixed. The disadvantages are a lack of geographic detail, the institutionalized population is not sampled, and variables are primarily limited to social and demographic factors. This data source links death certificate information to eight samples from the Current Population Survey (CPS) (Rogot, Sorlie, and Johnson 1992a; Rogot, Sorlie, Johnson, and Schmitt 1992b; National Heart, Lung, and Blood Institute 1992).

2.2. *Details on socio-economic variables available in the data sources*

Each of the data sources collects data on socio-economic variables (Table 1). The questions on education are as follows:

DECEDENT'S EDUCATION (Specify only highest grade completed):

Elementary/Secondary (0–12) College (1–4 or 5+)

(National Vital Statistics Mortality Data File (VSUS))

What was the highest grade or year of regular school the person ever completed? (Mark (x) only one box)

☐ Less than 5 years ☐ 5–7 years ☐ 8 years ☐ 9–11 years
☐ High school graduate ☐ 1–3 years of college ☐ 4 years of college or more

(National Mortality Followback Survey (NMFS))

What is the highest grade or year of regular school _____ has ever attended? (Circle appropriate number) ☐ Never attended or kindergarten only

Elementary . . . 1 2 3 4 5 6 7 8

High school . . . 9 10 11 12

College . . . 1 2 3 4 5+

Did _____ finish (number from previous question) (grade/year)? ☐ Y ☐ N

(NHANES Epidemiologic Followup Study (NHEFS))

What is the highest grade or year of regular school _____ has ever attended?

Did _____ complete that grade?

(U.S. National Longitudinal Mortality Study (NLMS))

In the studies with questions about "regular school," the term refers to education that leads toward an elementary or high school diploma or a college, university, or professional degree. These include public and private and day and night schools (Rogot et al. 1992b).

The questions on occupation and industry are:

DECEDENT'S USUAL OCCUPATION

(Give kind of work done during most of working life. Do *not* use retired.)

KIND OF BUSINESS/INDUSTRY

(National Vital Statistics Mortality Data File (VSUS))

Did the person EVER work at a paying job or a business full or part time?

☐ Yes ☐ No

Of all the PAID jobs or businesses the person ever had, what KIND OF WORK did he or she do the longest? (For example, electrical engineer, stock clerk, typist, farmer, in Armed Forces, etc.)

For how many years did the person do this kind of work?

☐ Less than one year ☐ 1 to less than 5 years ☐ 5 to less than 10 years

☐ 10 to less than 20 years ☐ 20 to less than 30 years ☐ 30 to less than 40 years

☐ 40 years or more

In this occupation, what KIND OF BUSINESS OR INDUSTRY did he or she work in the longest? Describe the activity at the location where employed. (For example: TV and radio manufacturing, retail shoe store, State Labor Department, farm, Armed Forces, etc.)

Was the person employed at a paying job or business up until the time he or she died?

☐ Yes ☐ No

(National Mortality Followback Survey (NMFS))

During the past 2 weeks, did _____ work at any time at a job or business, not counting work around the house?

☐ Y ☐ N

INCLUDE UNPAID WORK IN THE FAMILY FARM OR BUSINESS.

Even though _____ did not work during those 2 weeks, did _____ have a job or business?

☐ Y ☐ N

Was _____ looking for work or on layoff from a job?

☐ Y ☐ N

Which, looking for work or on layoff from a job?

☐ looking ☐ layoff ☐ both

For whom did _____ work?

ENTER NAME OF COMPANY, BUSINESS, ORGANIZATION, OR OTHER EMPLOYER.

For whom did _____ work at – *last* full-time civilian job or business lasting 2 consecutive weeks or more?

ENTER NAME OF COMPANY, BUSINESS, ORGANIZATION, OR OTHER EMPLOYER.

What kind of business or industry is this? (For example, TV and radio, manufacturing, retail shoe store, State Labor Department, farm.)

What kind of work was _____ doing? (For example, electrical engineer, stock clerk, typist, farmer.)

What were _____'s *most important* activities or duties at that job? (For example, types, keeps account books, files, sells cars, operates printing press, finishes concrete.)

From the answers to these questions, the interviewer completes the following:

Was _____ an employee of a *private* company, business or individual for wages, salary or commission? . . . P

– a *Federal* government employee? ... F

– a *State* government employee? ... S

– a *Local* government employee? ... L

– self-employed in *own* business, professional practice, or farm? IF NOT FARM, ASK: Is the business incorporated?

Yes ... I

No ... SE

– working *without pay* in business or farm? ... WP

– *never worked* or never worked at a full-time civilian job lasting 2 weeks or more ... NEV

(NHANES Epidemiologic Followup Study (NHEFS))

The questions on income are as follows:

What was the family's income in 1985? (Please include the person's income and the income of all other related persons living in the same household as the person. Include money from jobs, social security, retirement income, unemployment payments, public assistance, etc. Also include income from interest, dividends, net income from business, farm, or rent, and any other money income received.)

☐ Less than $5,000 ☐ $5,000–$6,999 ☐ $7,000–$8,999 ☐ $9,000–10,999

☐ $11,000–$12,999 ☐ $13,000–$14,999 ☐ $15,000–$16,999

☐ $17,000–$18,999 ☐ $19,000–$21,999 ☐ $22,000–$24,999

☐ $25,000 and over

(National Mortality Followback Survey (NMFS))

Including wages, salaries, self-employment, and any other source of income we just talked about, was the *total combined family income* during the *last 12 months* – (that is, yours, READ NAMES OF ALL FAMILY MEMBERS, INCLUDING ARMED FORCES MEMBERS LIVING AT HOME) – more or less than $20,000?

☐ Less than $20,000 ☐ $20,000 or more ☐ No income

Of those income groups, which letter best represents the *total combined family income* during the *last 12 months* (that is, yours, READ NAMES OF ALL FAMILY MEMBERS, INCLUDING ARMED FORCES MEMBERS LIVING AT HOME)? Include all sources of income we just talked about.

[card shows $1000 increments up to $20,000, $5,000 increments up to $50,000, $10,000 increments to $80,000 and $80,000 and over.]

Now, please think about your *family income* during (NAME OF LAST MONTH). Which letter best represents the *total combined family income* during (NAME OF LAST MONTH) (that is, yours, READ NAMES OF ALL FAMILY MEMBERS, INCLUDING ARMED FORCES MEMBERS LIVING AT HOME)? Again, include all sources of income we just talked about.

(NHANES Epidemiologic Followup Study (NHEFS))

3. Illustrative Results

Death rates were calculated for each data set to show the association of socio-economic factors with mortality for persons aged 25–74 years. Tables 2 and 3 present illustrative results of information available in these data sources using comparable measures. In Table 2, only bivariate relationships are shown; in Table 3, the effect of age is controlled. The following section highlights methodological differences in calculating death rates for the different data sets.

3.1. Differences in calculating death rates

Preliminary estimates of mortality by level of education for the National Vital Statistics Mortality Data File (VSUS) are based on vital statistics for 21 of the 50 states for 1989. Deaths for which level of education was not known were proportionally allocated for each age–race–sex group according to the proportion of deaths with known educational level for each age–race–sex group. The population estimate used to compute age-race–sex–education-specific death rates for 1989 employed education data from the Current Population Survey (CPS) and the 1990 Decennial Census. The total 1990 census population for each age–race–sex group for the 21 states was multiplied by the corresponding age–race–sex-education proportions for the entire United States from the 1989 CPS (Rosenberg et al. 1992).

For calculating death rates with the 1986 National Mortality Followback Survey (NMFS), decedents for whom level of education was not stated were assigned an education level according to the proportion of records with stated level of education. The 1986 National Health Interview Survey (NHIS) provided the denominator data for the death rates presented in Tables 2 and 3.

Published data on the NHANES Epidemiological Followup Study (NHEFS) are used to calculate the death rates shown in Tables 2 and 3 (Feldman et al. 1989; Makuc, Feldman, Kleinman, and Pierre 1990). To calculate annual death rates in the NHEFS sample, the following formula is used

$$\text{Death Rate} = [D_{71-84}/[[S_{84} + (D_{71-84})/2] * 10 \text{ years}]] * 100,000 \quad (1)$$

where D is the number of deaths during the 10-year follow-up and S is the number of survivors at the end of the 10-year follow-up period. Deaths are assumed to occur uniformly throughout the period rather than being clustered in the first part of the period.

The U.S. National Longitudinal Mortality Study (NLMS) decedents for whom

Table 2. Death rates for persons aged 25–74 by sex, age, race, marital status, education: Comparison of data sources

[Data Sources: National Vital Statistics Mortality Data File (VSUS), 1989; U.S. National Mortality Followback Survey (NMFS), 1986; NHANES Epidemiologic Followup Study (NHEFS), 1971–1984; U.S. National Longitudinal Mortality Study (NLMS), 1979–1985. Rates per 100,000 population.]

Characteristic	Males				Females			
	VSUS	NMFS	NHEFS	NLMS	VSUS	NMFS	NHEFS	NLMS
Age at death								
25–34 years	196	198	–	138	71	70	–	67
35–44	285	293	–	286	129	140	–	175
45–54	573	652	–	780	326	362	–	411
55–64	1511	1649	1940	1982	858	906	991	1006
65–74	3278	3729	3687	4407	1886	2120	2508	2234
Race								
White	801	902	–	1038	485	533	–	585
Black	1275	1392	–	1464	722	762	–	791
Marital status								
Married[@]	765	790	–	1018	371	393	–	456
Widowed	3711	4156	–	4156	1794	1895	–	1783
Divorced	1249	1708	–	1249	555	574	–	554
Never married	763	989	–	786	374	424	–	475
Family income								
<$9,000	***	3443	–	2390[#]	***	1536	–	1161[#]
$9,000–18,999	***	1705	–	1030	***	720	–	517
$19,000–24,999	***	712	–	627	***	456	–	317
$25,000+	***	380	–	557	***	241	–	296

Education								
0–8 years	1484	2195	–	2442	897	1169	–	1334
9–11	1123	1696	–	1540	587	822	–	800
12	883	819	–	863	533	532	–	469
13–15	656	621	–	655	391	345	–	398
16 or more	448	406	–	465	249	243	–	301

[@] Includes separated.

[#] The income categories in the NLMS are as follows: <$10,000, 10,000–19,999, 20,000–24,999, and 25,000+

*** Not available

– Not calculated

Note: Figures for NHEFS were calculated from data presented in Makuc et al. (1990)

Table 3. Death rates for persons aged 25–74 by age, sex, and education: Comparison of data sources
[Data Sources: National Vital Statistics Mortality Data File (VSUS), 1989; U.S. National Mortality Followback Survey (NMFS), 1986; NHANES Epidemiologic Followup Study (NHEFS), 1971–1984; U.S. National Longitudinal Mortality Study (NLMS), 1979–1985. Rates per 100,000 population.]

Age at death and educational attainment	Males				Females			
	VSUS	NMFS	NHEFS	NLMS	VSUS	NMFS	NHEFS	NLMS
Age-adjusted								
<12 years	882	1280	–	1217	431	591	–	615
12	870	789	–	895	451	458	–	450
13+	638	593	–	709	384	354	–	409
Age-specific								
25–34								
<12 years	359	497	–	234	130	179	–	115
12	220	193	–	164	75	59	–	62
13+	123	109	–	94	50	48	–	56
35–44								
<12 years	454	644	–	487	187	297	–	247
12	341	326	–	246	135	127	–	164
13+	209	178	–	222	106	99	–	140
45–54								
<12 years	713	1158	–	1042	363	519	–	550
12	632	574	–	756	341	356	–	366
13+	450	443	–	555	282	257	–	325
55–64								
<12 years	1673	2251	1823	2469	826	1114	1105	1236
12	1668	1427	3305	1796	912	866	853	878
13+	1213	1210	–	1479	800	695	–	846

65–74								
<12 years	2964	4421	3854	4959	1655	2120	2696	2470
12	3820	3584	3249	4102	2121	2314	2085	2059
13+	3096	2746	–	3269	1856	1779	–	1847

– Not calculated

Note: 1. Figures for NHEFS come from Feldman et al. (1989). The education categories for the NHEFS are <12 and 12+ years of education.

2. U.S. Standard Million Population (relative distribution of 1940 enumerated population of U.S. totaling 1,000,000 used as standard for calculating age-adjusted rates.

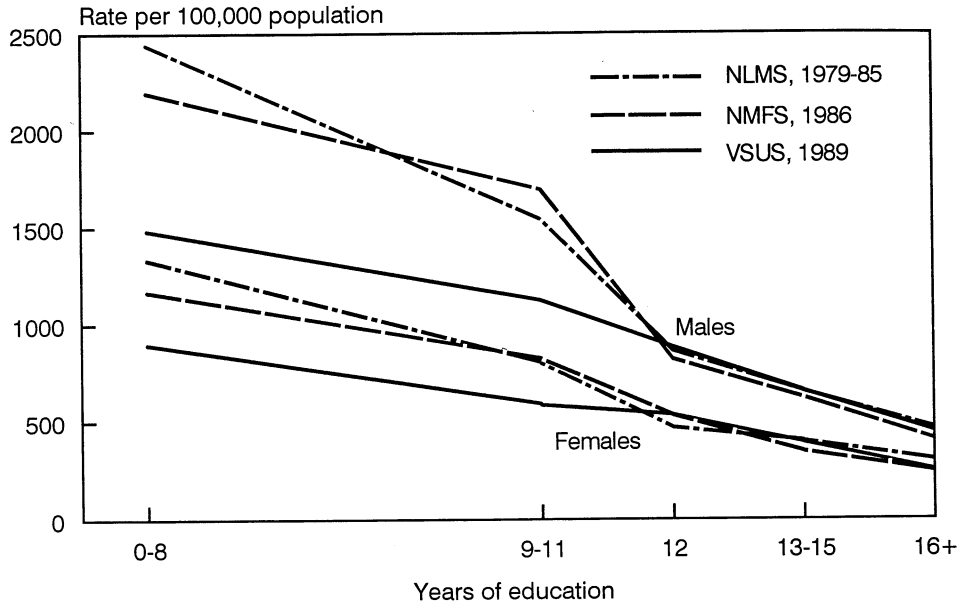

Fig. 7. Comparison of death rates in three data sources by education: United States, selected years 1979–89

education was not reported were assigned an educational level according to the proportion of records with stated level of education. The death rate here is defined as the number of deaths divided by person-years lived, where person-years lived is a measure of exposure to the risk of death. To calculate annual death rates in the NLMS sample, the following formula is used

$$\text{Death Rate} = [D_{80-85}/[[S_{85} + (D_{80-85})/2] * 5 \text{ years}]] * 100,000 \quad (2)$$

where D is the number of deaths during the five-year follow-up and S is the number of survivors at the end of the five-year follow-up period. Deaths are assumed to occur uniformly throughout the period.

3.2. *Selected results*

Table 2 shows death rates by age, race, marital status, and education for the National Vital Statistics Mortality Data File (VSUS), National Mortality Followback Survey (NMFS), U.S. National Longitudinal Mortality Study (NLMS), and to a limited extent, NHANES Epidemiologic Followup Study (NHEFS). The associations shown in Table 2 do not control for any other characteristics.

As expected, each of the data sets indicates monotonically increasing mortality by age for both males and females. Death rates calculated from the various data sources are about the same for those dying at younger ages but begin to diverge at older ages. For the age group 65–74 years, the NLMS rate exceeded the VSUS rate by 34 percentage points for males, whereas the NHEFS rate exceeds the VSUS rate by 33 percentage points for females. This, in part, reflects that age is measured at the baseline in the longitudinal studies (NLMS and NHEFS), while it represents age at death in the

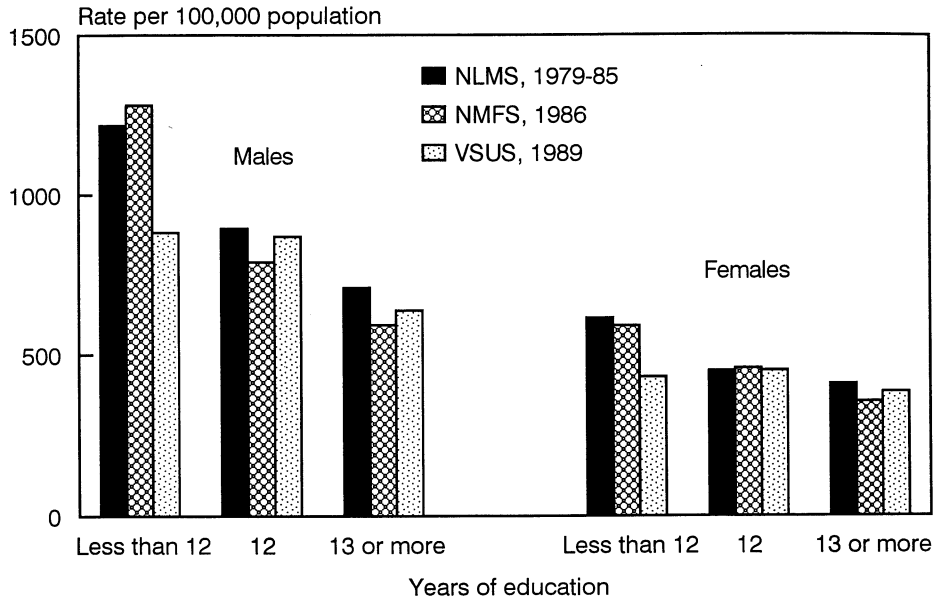

Fig. 8. Comparison of age-adjusted death rates in three data sources by education: United States, selected years 1979–89

cross-sectional studies (VSUS and NMFS). Another source of differences may relate to the quality of data on age in different data sources, particularly at older ages. In the United States, Social Security records are thought to be the most accurate source of data on extreme age because proof of age has been required when enrolling in Social Security benefit programs since 1965. For the older age groups included in this paper, the percentage agreement on exact age between matched Social Security records and death certificates for persons 65–69 years was 95.4 and for persons 70–74 was 94.6; the percentage agreement within five-year intervals was 98.1 and 97.9, respectively (Kestenbaum 1992).

Death rates for the Black population are consistently higher than for the white population in each data set. For males, the mortality race ratio is between 1.4 and 1.6 in VSUS, NMFS, and NLMS. For females, the corresponding ratios are between 1.4 and 1.5 for each of the data sets.

Marital status differentials in mortality were not completely consistent across the data sets (Table 2), but in each data source, death rates were highest for widows and widowers followed by divorced persons. Rates were low for never married and married persons in each data set; however, neither was consistently lowest across the data sets. Given the tremendous differences in marital status by age, inconsistencies, to some degree, may reflect age differences in the data sets. Additionally, marital status in the NLMS refers to the baseline survey and the rates for married persons may be higher in this data source because this data item is collected in the baseline survey and would include persons widowed or divorced at the time of death.

The general mortality pattern associated with family income is consistent across the data sources (Table 2). In each data source regardless of sex, mortality levels decline

with increasing levels of education (Figure 7). Death rates calculated from the different data sets are about the same for those dying with 12 or more years of education, but vary for those with less than 12 years of education. For those with 0–8 years of education, the NLMS death rate exceeded the VSUS rate by 65 percentage points for males and by 49 percentage points for females. Differences in rates for the less educated suggest that there may be reporting errors in proxy reports for decedents with very little education (Shai and Rosenwaike 1989; Buescher and Leiss 1994).

Table 3 presents age-adjusted and age-specific death rates by education for both males and females. For each of the age groups shown for males in Table 3, mortality decreased monotonically with increasing education except for those aged 55–64 years in the NHEFS and 65–74 years in VSUS. Similarly, mortality decreased with increasing education at each age for females across each of the data sets with three exceptions: females aged 55–64 years in VSUS and those aged 65–74 years in VSUS and NMFS.

The age-adjusted mortality pattern was consistent across data sets for males but not for females (Table 3 and Figure 8). Age-adjusted rates indicated a monotonically inverse relationship between mortality and education for males in all three data sets. For females, this pattern held in only the NMFS and NLMS. In the VSUS, the age-adjusted rates for females with 12 years of education were about the same for females with less than 12 years of education.

4. Discussion

In this paper, several U.S. data sets useful for investigating socio-economic differentials in mortality are discussed. These data sets differ with respect to the variables available for analysis, data collection methodologies, and range of data analysis possible. The general patterns of association between measures of socio-economic measures and the risk of mortality are consistent across each of the data sets, although the absolute levels of mortality by specific covariates vary.

Cross-sectional data such as those derived from the National Vital Statistics Mortality Data File and the National Mortality Followback Survey are useful for descriptive purposes and for illuminating many relationships but are limited as a source of data needed to address causality and processes occurring over time. Routinely collected period data provide information on aggregate changes over time, but do not directly address individual changes over time. In the case of the National Mortality Followback Survey, retrospective questions on events in the last year of life and earlier, provide some data on processes leading towards death. However, longitudinal studies can provide much richer information on processes over time (the NHANES Epidemiologic Followup Study; Social Security Administration/Current Population Survey Matched Files, and the U.S. National Longitudinal Mortality Study). Data on changes in life circumstances, risk factors, health impairments, and disease processes enrich knowledge of mortality issues.

The data sources described in this paper include data collected using administrative records with no linkage to other data sources (National Vital Statistics mortality data file). This data collection methodology is a cost effective way to get information on

rare causes of death, small geographic areas, and complete description of the population of decedents. Three data sources (National Mortality Followback Survey, Social Security Administration/Current Population Survey Matched Files, and U.S. National Longitudinal Mortality Study) combine data collected using surveys with administrative records. This methodology combines cost effective savings of administrative records with a wider range of questions available from an interview. Combining the different data records (administrative record with interviews) also provides additional information for validation including multiple record sources and proxy and self-reports as well as some data on individuals at different points in time. Lastly, the NHANES Epidemiologic Followup Study collects data at multiple points in time using a variety of data collection methods. This complex design provides a tremendously rich data source with multiple sources of information for selected items. Limitations include a small sample size of decedents limiting analyses by cause of death or geographic area and missing data for some of the methods or at certain points in time.

Each of these data sets provides an opportunity to further our knowledge of socio-economic differentials in mortality in the United States, and will undoubtedly encourage more sophisticated work in the areas than has yet been done in the United States. Likewise, the practice of linking death certificate data to studies using other data collecting methodologies is expanding. The availability of data sets like those presented in this paper is very exciting, especially in view of the paucity of micro-data on socio-economic characteristics of decedents in the U.S. in the past. Together, these data sets and new variables contained therein provide considerable potential for understanding the socio-economic correlates of mortality in the United States.

5. References

Ashley, J.S., Cole, S.K., Kilbane, M.P.J. (1991). Health Information Resources: U.K. – Health and Social Factors. In Oxford Textbook of Public Health, 2nd ed., Vol. 2: Methods of Public Health, eds. W. W. Holland, R. Detels, and G. Knox, Oxford: Oxford University Press.

Babbie, E. (1986). The Practice of Social Research. Belmont, CA.: Wadsworth Publishing Co.

Brackstone, G.J. (1989). Discussion of Administrative Records in Censuses. Proceedings of the U.S. Bureau of the Census: Fifth Annual Research Conference, Arlington, VA.

Buescher, P.B. and Leiss, J.K. (1994). Race, Education, and Mortality in North Carolina. CHES Studies, Raleigh, N.C.: The State Center for Health and Environmental Statistics.

Caselli, G. (1991). The Quest for an Interpretation of Mortality Differences: Socio-Economic Background and Adult Mortality. In Socio-Economic Differential Mortality in Industrialized Societies, Vol. 7. Paris: INED, INSEE, and CICRED, 229–241.

Desplanques, G. (1976). La Mortalité des Adultes Suivant le Milieu Social 1955–71. Institute National de la Statistique et des Etudes Economique, 195, series D, no. 44.

Duleep, H.O. (1986). Measuring the Effect of Income on Adult Mortality Using Longitudinal Administrative Record Data. The Journal of Human Resources, XXI, 238–251.

Duleep, H.O. (1989). Measuring Socioeconomic Mortality Differentials Over Time. Demography, 26, 345–351.

Eriksen, R. (1983). Socio-Economic Classification in Swedish Statistics. Paper presented at the Third Meeting of the UN/WHO/CICRED Network on Socio-Economic Differential Mortality. Rome, Italy.

Feldman, J.J., Makuc, D.M., Kleinman, J.C., and Cornoni-Huntley, J. (1989). National Trends in Educational Differentials in Mortality. American Journal of Epidemiology, 129, 919–933.

Fox, A.J., Goldblatt, P.O., and Adelstein, A.M. (1992). Selection and Mortality Differentials. Journal of Epidemiology and Community Health, 36, 69–79.

Gittelsohn, A. and Royston, P.N. (1982). Annotated Bibliography of Cause-of-Death Validation Studies, 1958–80. National Center for Health Statistics. Vital Health Statistics 2(89).

Groves, R.M. (1989). Survey Errors and Survey Costs. New York: John Wiley and Sons.

Kestenbaum, B. (1992). A Description of the Extreme Aged Population Based on Improved Medicare Enrollment Data. Demography, 29, 565–580.

Kovar, M.G. (1989). Data Systems of the National Center for Health Statistics. National Center for Health Statistics. Vital Health Statistics, 1(23).

Kristofersen, L. (1979). Occupational Mortality. Report no. 19, Statistisk sentralbyrå. Oslo, Norway.

Kunst, A.E. and Mackenbach, J.P. (1994). The Size of Mortality Differences Associated with Educational Levels in Nine Industrialized Countries. American Journal of Public Health, 84, 932–937.

Makuc, D.M., Feldman, J.J., Kleinman, J.C., and Pierre, Jr, M.B. (1990). Sociodemographic Differentials in Mortality. In Health Status and Well-Being of the Elderly, eds. J.C. Cornoni-Huntley, R.R. Huntley, J.J. Feldman, Oxford: Oxford University Press, 155–171.

National Center for Health Statistics (NCHS). (1989). Report of the Workshop on Improving Cause-of-Death Statistics. Hyattsville, MD: Public Health Service.

National Center for Health Statistics (NCHS). (1990). National Death Index User's Manual. Hyattsville, MD: Public Health Service.

National Center for Health Statistics (NCHS). (1991). Vital Statistics of the United States, 1988, vol II, Mortality, Part A. Washington: Public Health Service.

National Center for Health Statistics (NCHS). (1993a). Healthy People 2000 Review. Health United States, 1992. Hyattsville, MD: Public Health Service.

National Center for Health Statistics (NCHS). (1993b). Advance Report of Final Mortality Statistics, 1991. Monthly Vital Statistics Report, Vol. 42, no. 2 suppl. Hyattsville, MD: Public Health Service.

National Center for Health Statistics (NCHS). (1993c). Annual Summary of Births, Marriages, Divorces, and Deaths: United States, 1992. Monthly Vital Statistics Report, Vol. 41, no. 13. Hyattsville, MD: Public Health Service.

National Center for Health Statistics (NCHS). (1994). Catalog of Electronic Data Products. Hyattsville, MD: Department of Health and Human Services.

National Heart, Lung, and Blood Institute (NHLBI). (1992). NLMS Public Use Data File, 1979–1985 Follow-up. Bethesda, MD: National Institutes of Health.

Pappas, G., Queen, S., Hadden, W., and Fisher, G. (1993). The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986. New England Journal of Medicine, 329: 103–109.

Powell-Griner, E. and Rosenberg, H.M. (1991). New Data on Socio-Economic Differential Mortality in the United States. Socio-Economic Differential Mortality in Industrialized Societies. INED: Paris, 182–199.

Robey, B. (1989). Two Hundred Years and Counting: The 1990 Census. Population Bulletin, 44.

Rogers, R.G. (1992). Living and Dying in the U.S.A.: Sociodemographic Determinants of Death Among Blacks and Whites. Demography, 29, 287–303.

Rogot, E., Sorlie, P.D., and Johnson, N.J. (1992a). Life Expectancy by Employment Status, Income, and Education in the National Longitudinal Mortality Study. Public Health Reports, 107, 457–481.

Rogot, E., Sorlie, P.D., Johnson, N.J., and Schmitt, C. (1992b). A Mortality Study of 1.3 Million Persons by Demographic, Social, and Economic Factors: 1979–1985 Follow-up. Bethesda, MD: National Institutes of Health.

Rosenberg, H.M., Burnett, C., Maurer, J.D., and Spirtas, R. (1993). Mortality by Occupation, Industry, and Cause of Death: 12 Reporting States, 1984. Monthly Vital Statistics Report. Vol. 42, no. 4, suppl. Hyattsville, MD: National Center for Health Statistics.

Rosenberg, H.M., Makuc, D.M., Hoyert, D.L., Maurer, J.D., Singh, G.K., and Feldman, J.J. (1992). Race and Educational Attainment in Mortality Statistics. Paper presented at the American Public Health Association. Washington, D.C.

Seeman, I., Poe, G.S., and Powell-Griner, E. (1993). Development, Methods, and Response Characteristics of the 1986 National Mortality Followback Survey. National Center for Health Statistics. Vital Health Statistics, Series 1, no. 29.

Shai, D. and Rosenwaike, I. (1989). Errors in Reporting Education on the Death Certificate: Some Findings for Older Male Decedents from New York State and Utah. American Journal of Epidemiology, 130, 188–192.

Sorlie, P.D., Backlund, E., and Keller, J.B. (1995). U.S. Mortality by Economic, Demographic, and Social Characteristics: The National Longitudinal Mortality Study. American Journal of Public Health, 85, 949–956.

Sterling, T., Rosenbaum, W., and Weinkam, J. (1993). Income, Race, and Mortality. Journal of the National Medical Association, 85, 906–911.

U.S. Bureau of the Census. (1978). The Current Population Survey: Design and Methodology. Technical Paper 40. Washington, DC.: Department of Commerce.

U.S. Department of Health and Human Services. (1991). Healthy People 2000: National Health Promotion and Disease Prevention Objectives. Washington, DC.: Public Health Service.

Valkonen, T. (1982). Socioeconomic Mortality Differentials in Finland. Paper presented at the conference Health for All by the Year 2000 in the Nordic Countries, Gothenburg, Sweden.

Valkonen, T. (1993). Problems in the Measurement and International Comparisons of Socio-economic Differences in Mortality. Social Science Medicine, 36, 409–418.

Valkonen, T., Javanainen, M., Koskinen, S., Martelin, T., and Martikainen, P. (1991). Research on Socio-economic Differential Mortality at the Department of Sociology, University of Helsinki. In Socio-Economic Differential Mortality in Industrialized Societies. Vol 7, Paris: INED, INSEE, and CICRED, 55–81.

World Health Organization. (1977). Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, Based on the Recommendations of the Ninth Revision Conference, 1975. Geneva: World Health Organization.

Received January 1994
Revised July 1995