COMPARABILITY BETWEEN ITEMS REPORTED ON THE DEATH CERTIFICATE AND INFORMANTS ON THE 1993 NATIONAL MORTALITY FOLLOWBACK SURVEY

Susan G. Queen, National Center for Health Statistics
6525 Belcrest Road, Hyattsville, Maryland 20782

KEY WORDS: Comparability, informants, death certificate, agreement.

The major source of information on annual mortality data in the United States is the standard death certificate. These vital records provide statistical information which is used to evaluate and design policies and programs in a variety of agencies and State and local health departments. The accuracy of the standard certificates, therefore, is critical when considering the varied uses of vital records. Items reported on death certificates provide the basis for national and State mortality statistics. Information from death certificates has been a major source to study the general health of the U.S. population and to identify many of our public health problems. This information is often utilized in determining policy and plans to allocate medical and health care services, and to evaluate the success of programs which have been designed to address these problems. The death certificate can be introduced in court as a legal document, to apply for insurance benefits, and in the transfer of title of property. For these varied reasons it is crucial to strive for accuracy in the registration of this information (Rosenberg et al. 1979; Davis 1988; Rosenberg 1989).

Using death certificate information and supplemental data from informants, this report will explore the quality of the data that is provided on the death certificate. There is no way to determine which source of data is more accurate, the death certificate or the informant survey, particularly since the same informant provided the information (in most cases). For this reason, the death certificate will be used as the more reliable source, since the information was collected immediately after the death and the informant survey had a significantly greater interval between the death and data collection.

Data Sources

The 1993 National Mortality Followback Survey (NMFS) will supplement information from death certificates in the vital statistics file with information on important characteristics of the decedent. These include use of health services during the last year of life, socioeconomic status, aspects of life style and health behaviors, and other factors that may be related to how and when death occurs. Supplemental information is obtained from informants listed on death certificates filed in State and independent registration areas. Information is also collected from health care facilities utilized in the last year of life, as well as from medical examiners/coroners, when applicable. The sample of approximately 20,000 deaths will be weighted to nationally representative estimates using post-stratified ratio estimation. The NMFS will sample deaths to individuals aged 15 and over who died in the United States in 1993. The 1993 NMFS will place special emphasis on young and external causes of death, so that we may better understand factors that may be related to these types of deaths (USDHHS 1990). The sampling frame will consist of death certificates selected from the Current Mortality Sample (CMS), a 10 percent sample of the States' death certificates received by NCHS about three months after they occur. Data from the informants will be collected by telephone and personal interviews. Facilities will receive a mailed questionnaire, and medical examiners/coroners will provide autopsy reports, toxicology findings, and investigative reports. The data collection agency for the 1993 NMFS is the U.S. Bureau of the Census.

Data from the following sources will be collected for the 1993 NMFS:

1. Death Certificates: socio-demographic characteristics and multiple cause ICD-9 codes.
2. Informant questionnaires: data on access and utilization of health services, sources of payment for health care, functional limitations, cognitive functioning, socioeconomic characteristics, health and life-style behaviors (alcohol and tobacco use, for example), life events, medical conditions, and medical devices.
3. Facility Abstract Records (FARS): data on length of stay, type of facility, diagnostic and surgical procedures, and billing information.
4. Medical Examiner/Coroner records: autopsy reports, toxicology findings, and investigative reports.

The pretest to the 1993 NMFS was conducted in four states in the four major Census regions. These states also provide examples of the various types of systems that will be encountered in the collection of medical examiner/coroner data. (These systems vary widely by state, jurisdiction, degree of centralization, etc.) The informant phase of the pretest has recently been completed, and this paper will present some

findings from that pretest and discuss implications for the main survey. This paper will focus on the comparability of items reported on death certificates with items reported on the informant questionnaire.

Methods

A sample of 807 individuals aged 15 years or older who died during the first quarter of 1992 were selected from the CMS for the pretest. Since the main survey will oversample external causes which occur in disproportionate numbers to the young, the pretest oversampled deaths to younger persons in order to determine the level of difficulty involved in oversampling external causes. The pretest achieved an overall response rate of 80.8% (n=652). Approximately 200 deaths per state were sampled. Of the total sample cases, 11.6% refused to participate, and the remainder of the nonresponse was related to problems in locating the informant. The variables that will be examined for comparability between the death certificate and the informant survey are: age, gender, race, hispanic origin, marital status, veteran status, and education. Occupation was to have been included in the analysis, but was not one of the variables coded from the death certificate, thereby excluding the opportunity to study occupation/industry agreement between the data sources. However, the addition of education as a variable is new to the death certificate, and the NMFS provides a unique opportunity to examine reporting of this information. Education is a useful measure of socioeconomic status, which has been shown to be closely associated with mortality. For some variables, the informant survey had different response categories than did the death certificate. In these cases (education, marital status, race, and hispanic origin) the informant categories were recoded to match the death certificate codes.

The way in which education was measured on the death certificate differed from the informant survey. The death certificate asks for decedent's education, specified by highest grade **completed**. The informant survey asks for the highest grade **attended**, with a follow-up question asking if the decedent completed the grade. There is a subtle difference, therefore, in the way that education is collected.

Although an Hispanic identifier has been added to the death certificate in order to collect more detailed data on Hispanic mortality, this analysis will not look at the detailed codes. The pretest had too few cases in these categories to do meaningful analysis, however, the main survey will provide enough cases to examine the detailed Hispanic identifier. Hispanic origin as a dichotomous 'yes/no' variable from the pretest will be included here.

Because this analysis is on pretest data, national estimates are not possible: these deaths are not representative of the U.S. The purpose of this paper is exploratory in nature, to determine the extent to which the informant survey is accurately reflecting information given by the same informant regarding decedent characteristics. The findings here may have implications for the main survey regarding the wording of questions and their response categories.

Analysis

In the comparison between the two sources of data, the death certificate is being used as the standard for the comparison, and the informant survey will be compared against this standard. Percent of agreement is based on the number of cases for which there is a valid response on both the death certificate and the informant questionnaire. Cases for which there are missing values on any of the variables are excluded from the analysis. Potential sources of error are: coding errors on the informant survey, problems of recall for the informant survey, and reporting errors on the death certificate. Agreement will be examined by informant type (spouse, parent, child, etc.) and by whether the informant for the pretest was the original informant listed on the death certificate.

Overall agreement between the items on the death certificate and on the informant survey is presented in Table 1. Age, gender, race, Hispanic origin, and marital status all have fairly high agreement between the two sources of data. Veteran status has only 75.5% agreement and education is quite low with only 47.9% agreement. In this table education was compared using single years of completed education. In the following tables, education was collapsed into four categories: less than high school, high school, 1 to 3 years of college, and 4 or more years of college.

When education is collapsed into the four categories mentioned above, the overall agreement between the death certificate and the informant survey is 70.1%, considerably higher than when left in single years.

The informant survey collected information on the relationship of the informant to the decedent, whether the informant was the spouse, parent, child, etc. Table 2 presents the percent of agreement on education in 4 categories by type of informant. The highest agreement occurs when the informant is the parent of the decedent. The lowest agreement is

present when the informant is in the category 'other' which includes more distant relatives, friends, and staff (if the decedent died in an institution).

Agreement on education by age of decedent is presented in Table 3, in order to explore further the reason for the low agreement on this variable. The highest agreement (78.6%) occurs when the decedent is in the youngest age group, 15 to 24 years. The lowest agreement (68.4%) is when the decedent is aged 25 to 44. The youngest age group is most likely to have an parent as an informant, and as was demonstrated in Table 2, when the parent was the informant, the agreement between the data sources is higher.

Table 4 provides the percent of agreement for the decedent's veteran status by the type of informant. The greatest agreement occurs when the informant was the spouse (84.6%) and the lowest agreement is when the informant was in the 'other' category (66.6%). It is unfortunate that this 'other' category did not request the specifics of the relationship to the decedent, because approximately 20% of the informants were in this category. For the main survey, detailed information will be collected for this category.

Conclusions

Overall agreement between the death certificate and the informant survey was fairly high for most items. Veteran status had less agreement, and education had poor agreement. When education is coded into four meaningful categories, the agreement is considerably greater, but is still not as high as items such as gender or race. This poses a dilemma for the researcher who is relying upon this information as a reliable source of data. Some possible solutions are to consider the type of informant, since this item was related to comparability, and to examine the manner in which each state collected the information, since the method of reporting education was not identical for each state.

Limitations of the Data

A major limitation of the informant survey is the use of proxy or surrogate respondents to provide the desired information. This same limitation applies to the information collected on the death certificate, however. Epidemiologic research often makes use of surrogate interviews for information on cigarette smoking, dietary history, patterns of alcohol consumption, etc. Many of these studies have found that validity is dependent on: the topic of the question, the level of specificity requested, and the relationship of the surrogate to the individual (McLaughlin et al. 1987, Rogot and Reid 1976, Lerchen and Samet 1986). Fortunately, the types of items collected on the death certificate are those that are believed to be accurate when asked of a surrogate. For the most part, they require little historical information on the decedent, are not as subject to the issue of recall, and are not particularly detailed.

An issue of major concern is that the states do not collect information on education in a standard format. Although there is a standard death certificate, each state determines the exact format and wording of the items on the death certificate. While the pretest was limited to four states, only two of those states collected education identically. The other two had subtle differences in the wording and response categories. While these differences are minor, they may well contribute to lack of agreement between the informant survey and the death certificate.

The results reported here will be useful for the final revisions of the informant survey to the 1993 NMFS. If discrepancies exist between the death certificate and the informant survey, it may suggest that the informant survey be patterned more closely after the death certificate. Issues of survey design involving format, question wording, and response categories can be revised appropriately to improve correspondence between these two sources of data.

Bibliography

Gittelson, A. and P.N. Royston. (1980), Annotated Bibliography of Cause-of-Death Validation Studies: 1958-1980, Vital and Health Statistics, Series 2. No. 89. Washington, D.C., U.S. Government Printing Office.

Davis, H. (1988), "The Accuracy of Industry Data From Death Certificates for Workplace Homicide Victims," American Journal of Public Health, 78(12), 1579-1581.

Lerchen, Mary L. and Jonathan M. Samet. (1986), "An Assessment of the Validity of Questionnaire Responses Provided by a Surviving Spouse," American Journal of Epidemiology, 123(3), 481-489.

McLaughlin, Joseph K., Michael S. Dietz, Eric S. Mehl, and William J. Blot. (1987), "Reliability of Surrogate Information on Cigarette Smoking by Type of Informant," American Journal of Epidemiology, 126(1), 144-146.

Rogot, Eugene and D. D. Reid. (1976), "The Validity of Data from Next-of-Kin in Studies of Mortality Among Migrants," International Journal of Epidemiology, 4(1), 51-54.

Rosenberg, Harry M. (1989), The Nature and Accuracy of Cause-of-Death Data. Proceedings of the Workshop on Improving Cause-of-Death Statistics, Vital and Health Statistics, Series 4. Washington, D.C., U.S. Government Printing Office.

Rosenberg, H. M., D. Burnham, R. Spritas, V Valdisera. (1979), Information from the Death Certificate: Assessment of the Completeness of Reporting. In L. Delbene and R. Scheuren, Statistical Uses of Administrative Records With Emphasis on Mortality and Disability Research. Washington, D.C., U.S. Government Printing Office.

Schade, J. and G. M. Swanson. (1988), "Comparison of Death Certificate Occupation and Industry Data With Lifetime Occupational Histories Obtained by Interview: Variations in the Accuracy of Death Certificate Entries," American Journal of Industrial Medicine, 14(2), 121-136.

U.S. Department of Health and Human Services. (1990), Healthy People 2000: National Health Promotion and Disease Prevention Objectives. DHHS (PHS) 91-50212.

Table 1: Comparability Between Death Certificate and Informant

Variable	*Percent*	*N*
Age	88.2	618
Gender	99.3	642
Race	97.4	635
Hispanic Origin	97.8	629
Marital Status	94.1	635
Veteran Status	75.5	564
Education	47.9	542

Table 2: Education Collapsed to Four Groups by Informant Type

Informant Type	*Percent*	*N*
Spouse	67.8	112
Parent	75.8	148
Child	71.7	90
Sibling	71.1	81
Other	63.6	111

Table 3: Education Collapsed to Four Groups by Decedent Age

Decedent Age	*Percent*	*N*
15 to 24	78.6	66
25 to 44	68.4	130
45 to 64	73.4	238
65 or more	74.8	108

Table 4: Veteran Status by Informant Type

Informant Type	*Percent*	*N*
Spouse	84.6	130
Parent	74.3	148
Child	76.8	82
Sibling	75.3	81
Other	66.6	123