Death Certificate Statement of Occupation:
Its Usefulness in Comparing Mortalities

By ROBERT BUECHLEY, M.A., JOHN E. DUNN, Jr., M.D., M.S.P.H.,
GEORGE LINDEN, M.P.H., and LESTER BRESLOW, M.D., M.P.H.

SEVERAL STUDIES have indicated occupational factors in the occurrence of various chronic diseases. One may cite "chimney sweep's cancer" (1), lung cancer among chromate workers (2), and, more recently, higher rates for coronary heart disease (3) among London bus drivers as compared with bus conductors.

It would seem possible that a comprehensive survey of the frequency of other diseases among all occupations as recorded on death certificates, relating these to the numbers of persons in the occupations enumerated by the census, would provide valuable epidemiological leads—that is, the formulation of hypotheses about the occupational origin of certain diseases.

Such studies would require much detail and accuracy, both in the coding of cause of death and in the coding of occupation. After more than a half century of codification and revision,

Mr. Buechley is a social research technician, and Mr. Linden is a public health analyst, bureau of chronic diseases, California State Health Department. Dr. Breslow is chief of the bureau. Dr. Dunn is a field investigator, assigned to the bureau from the Field Investigations and Demonstrations Branch, National Cancer Institute, Public Health Service.

the International Statistical Classification of Diseases, Injuries, and Causes of Death has become a workable and consistent tool, with sufficient detail and sufficient accuracy for the purpose. This detail and accuracy has been achieved by querying causes of death that are ill defined or not understood and by a coding system that requires definite and unambiguous statements.

For the other axis of classification, occupation, two major coding systems have been developed in the United States since 1850: The Dictionary of Occupational Titles developed by and for the United States Employment Service and the system developed for the Bureau of the Census by Dr. Alba M. Edwards and others. These systems agree only in broad categories. Furthermore, the occupations reported on death certificates are not routinely coded and thus are not policed for completeness and accuracy as are reported causes of death. This has led to distrust of the accuracy of the occupation entries and to a natural reluctance to use the entries even in the search for epidemiological leads. Only by use, coding, and querying will occupation reporting increase in accuracy.

The only population base available by occupation for computation of rates is that provided by the decennial census enumeration. As we will point out, the definition for occupa-

Vol. 71, No. 11, November 1956   1105

tional classification on death certification and census enumeration are different. A major epidemiological consideration is the degree of agreement between these two entries in spite of the differences in definition.

Materials of the Study

The present study was undertaken by the bureau of chronic diseases, California State Health Department, to examine the errors resulting from the present, less than perfect, occupation reporting on death certificates. Though based on a chunk sample, the study offers evidence and conclusions which may be of some use. A recent investigation of lung cancer (4) made available life-long occupational data on certain persons. We took the opportunity to compare these data with death certificate statements of occupation for the same persons.

During 1949, 1950, and 1951, trained occupational analysts took complete occupational histories from 518 lung cancer patients and 518 control patients. The patients, ranging in age from 30 to 80 years, were located in university, county, private, and Federal hospitals. Patients in these hospitals came largely from the middle and working classes of the California population so that professional and managerial classes were under-represented. Since most of the hospitals were in urban areas, the farming population was also poorly represented.

The men in this group were followed through the death indexes of the California Department of Public Health for 1949 through 1952. Four hundred and twenty-three male deaths of all ages (mostly of the lung cancer cases) were found. A copy of the death certificate was obtained for each death and was attached to the occupational history form. Only the 304 males aged 35 to 64 were used for the balance of this study. Persons under 35 are often not "settled" in any occupation, and there are relatively few deaths from chronic disease occurring in this age range. Since persons 65 and over are frequently retired or have taken terminal occupations, only incomplete and less satisfactory population counts by occupation are available from the census enumeration for rate computations in this age group.

Assigning Occupational Categories

Before working out the problems in comparing death certificates and occupational histories when using closely specified occupations, it is necessary to know what is meant by occupation. Unfortunately, defining occupation is in itself a major problem. Besides the two coding systems noted, several assignment methods or definitions are used in various special fields. There is in addition the complete occupational history of the person concerned. For purposes of clarity, it seems desirable to list and explain the definitions and the five methods this study used for assigning individuals to occupation categories.

1. The entry of occupation on the death certificates often is the raw material in any study of death rates by occupations. Instructions in California and in most other States for entering occupations on death certificates are based on the model certificate proposed by the National Office of Vital Statistics of the Public Health Service. The actual entry, whether or not it is in accordance with these instructions, will be called "the death certificate occupation" or "the death certificate entry."

2. The complete occupational history of the person concerned is a chronological list and description of occupations and lengths of time spent at each. This information has been recorded for the limited group of persons in the present study. For the purposes of this investigation, it will be assumed that the listing is correct although some memory loss, and perhaps bias occurred. This list will be called "the occupational history."

3. The census method of assigning a person to an occupation is given in Bureau of the Census publications (5). It is an operational definition, derived from answers to the question, "What kind of work was he doing . . . at the job he held during the census week? . . . at the last job he held [if he were unemployed]?" In effect, then, the occupational populations derived from the census enumeration, to which the death certificate occupations need be related for the computing rates, are based on last occupation. This method will be called the census method, and the occupation derived from it will be called "the last occupation."

4. The National Office of Vital Statistics,

which recommends the rules for death certificate entries, defines the "usual occupation" of an individual as "the job he pursued for the longest part of his working life" (*6*). This will be called the NOVS definition, and the occupation derived from it will be called "the usual occupation."

5. Finally, on a trial basis, this study uses another method of assigning individuals to occupations. Epidemiological considerations require that the occupation be pursued for a sufficiently long time so that the exposure to the occupational environment can be expected to initiate pathological processes. There is no single answer to "how long is long enough?" but 5 years is offered as an estimate for chronic conditions. Thus, any jobs held for a total of 5 years or more will be called "the exposure occupations." Any one individual, therefore, can, according to this definition, appear in more than one occupational category.

All the occupations and identifying industries worked in for 5 years or more were abstracted from the occupational history onto a 3″ x 5″ card. This was accomplished by using the 3-digit numerical codes for occupation and industry in the Census Bureau's "Alphabetical Index of Occupations and Industries" (*7*).

Three classes of occupations were determined from the occupational history. These were (*a*) the usual occupation, determined by the NOVS definition, (*b*) the last occupation, determined by the census method, and (*c*) all the exposure occupations (frequently more than one for each individual), determined by the exposure method. These chosen occupations were also entered on the 3″ x 5″ card. Industry identification was used only as an aid to coding the occupation.

After the occupation codes had been entered from the occupational history, the occupation and industry from the death certificate were copied verbatim onto the card and then coded according to the 3-digit numerical code. If the code numbers for occupation agreed with respect to all 3 digits, the agreement between occupational history and death certificate entry was considered perfect. Agreement failed, however, when the death certificate and occupational history entries did not agree in respect to all 3 digits.

The number of cases having 3-digit agreement by each of the three methods is shown.

	Cases
All three methods	141
Usual occupation and last occupation but not exposure occupation	0
Usual occupation and exposure occupation but not last occupation	14
Last occupation and exposure occupation but not usual occupation	45
Usual occupation only	0
Last occupation only	26
Exposure occupation only	11
No agreement	67
Total number of cases	304

The percentage agreement for the "usual occupation" method is 51.0

$$\frac{(141+14=155)}{304} = \frac{155}{304}$$

For the "last occupation" method, it is 69.7

$$\frac{(141+45+26=212)}{304} = \frac{212}{304}$$

For the "exposure" method it is 69.4

$$\frac{(141+45+14+11=211)}{304} = \frac{211}{304}$$

The tabulation shows that there is a stable core of 141 persons (46.4 percent) with one and only one occupation determined on their occupational history, no matter which method is used. To this core are added special classes of persons to make up each total. Agreement in terms of usual occupation adds only 5 percent to this core. Last and exposure occupations show almost a 90 percent overlap, that is, the last occupation is usually held for 5 years. The highest percentage agreement is found with the last occupation method, a rather surprising result in view of the plain request for usual occupation on the death certificate.

Our finding indicates that to the people who furnish the information for entry on the death certificate "occupation" seems to mean current or present occupation. This finding appears to be true even if this last occupation is a short-time (less than 5 years) occupation, which tends to increase the number of last occupations reported correctly. The poor agreement of usual occupation apparently results from a tendency to ignore previous occupations if the last occupation has lasted several years.

Table 1. Degree of 3-digit agreement between death certificate statement of occupation and last occupation, in broad socioeconomic groups

3-digit occupations from death certificates		Socioeconomic groupings of 3-digit last occupations from occupational histories									
Socioeconomic group	Number of reported cases	Professional	Farmers and farm managers	Proprietors, managers, agents, and officials	Clerks	Sales	Crafts	Operatives	Service	Farm labor	Labor, excluding farm and mine
Professional	24										
Agree	16	16									
Disagree	8	2		1	1		3	1			
Farmers and farm managers	11										
Agree	5		5								
Disagree	6			1				2		2	1
Proprietors, managers, and officials	17										
Agree	13			13							
Disagree	4						3		1		
Clerks	14										
Agree	10				10						
Disagree	4						2		2		
Sales	22										
Agree	16					16					
Disagree	6		1	1	1		1	1	1		
Crafts	92										
Agree	69						69				
Disagree	23					1	5	8	3	1	5
Operatives	36										
Agree	22							22			
Disagree	14			1	1		4	4		2	2
Service	44										
Agree	36								36		
Disagree	8								7	1	
Farm labor	4										
Agree	3									3	
Disagree	1								1		
Labor, excluding farm and mine	33										
Agree	22										22
Disagree	11							5	4	1	1
Unknown and not in labor force	7										
Agree	0										
Disagree	7				1	1		3	1		1
Total	304	18	6	17	14	23	82	47	55	10	32
Agree	212	16	5	13	10	16	69	22	36	3	22
Disagree	92	2	1	4	4	7	13	25	19	7	10

Choosing the Method of Assignment

As indicated previously, the populations by occupation provided by census enumeration are based on last occupation. The validity of mortality rates by occupation, which are based on entries of occupation on death certificates on one hand and census enumerations on the other, is dependent on the degree to which death certificate entries represent last occupation. Fortunately for our purposes here, although in dis-regard of instructions, death certificate entries are most apt to be last occupation and not usual or some other prior occupation. The remainder of our presentation, therefore, will be concerned only with the agreement between last occupation revealed by occupational history and that entered on the individual's corresponding death certificate.

We attempted to group occupations according to a 2-digit code rather than the 3-digit

census code. This method resulted in 72 percent agreement on last occupation as against the 69.7 percent agreement with the 3-digit code. Further improvement in agreement, with loss of occupational precision, results from the use of 1-digit or social class code. This would give 76 percent agreement, as can be calculated from table 1. It does not appear that this increased agreement would be worth the loss of precision. In addition, these broad groupings are more socioeconomic than occupational, and thus exposures peculiar to specific occupations would be lost.

Another means by which agreement could be improved would be to code on environmental exposure, cutting across socioeconomic strata. This method would require groupings on the basis of similar exposure, such as grouping together "farmer" and "farm laborer," widely separated in the census system, or grouping together "aviator, army" and "aviator, commercial," also widely separated. Such a plan would entail development of a completely new coding system on exposure and would be entirely different from the approach with which this study

began. Hence, the 3-digit census codes are used as the basis of agreement or nonagreement for this study.

Accuracy of Reporting

Only a very few of the many individual occupations are present in sufficient numbers to be treated individually. Grouping these by broad socioeconomic groups, though admittedly not the ideal method, allows some estimates of the variability of reporting accuracy to be made.

Table 1 presents the death certificate occupation, as taken from the death certificate, in comparison with the last occupation, as taken from the occupational history. It attempts to give a concise report of the occupation shifts observed. Although occupations are grouped into the 10 socioeconomic groups used by the Bureau of the Census, it should be reemphasized that percent agreement is by detailed, 3-digit codes. The failures of agreement are of two kinds: upgrading and frank errors.

As evidence of errors of the first kind, the

Table 2. Comparisons between numbers designated as having a specific 3-digit occupation determined from occupational history, and the 3-digit occupation determined from death certificate entry

Occupational group	Correct number from occupational history	Number and percent designated on death certificate		Number and percent of death certificate occupations correct on a 3-digit basis		
		Number	Percent of correct number [1]	Number	Percent as designated on death certificate [2]	Percent of correct number from occupational history [3]
	(a)	(b)	(c)	(d)	(e)	(f)
Professional_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	18	24	133. 3	16	66. 7	88. 9
Farmers_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	6	11	183. 4	5	45. 5	83. 3
Managers_ _ _ _ _ _ _ _ _ _ _ _ _ _ _	17	17	100. 0	13	76. 5	76. 5
Clerks_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	14	14	100. 0	10	71. 4	71. 4
Sales_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	23	22	95. 6	16	72. 7	69. 6
Crafts_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	82	92	112. 2	69	75. 0	84. 1
Operatives_ _ _ _ _ _ _ _ _ _ _ _ _ _ _	47	36	76. 6	22	61. 1	46. 8
Service_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	55	44	80. 0	36	81. 8	65. 5
Farm labor_ _ _ _ _ _ _ _ _ _ _ _ _ _ _	10	4	40. 0	3	75. 0	30. 0
Other labor_ _ _ _ _ _ _ _ _ _ _ _ _ _	32	33	103. 2	22	66. 7	68. 8
Total_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	304	[4] 297	97. 7	212	69. 7	69. 7

[1] Column $b \div a \times 100$.
[2] Column $d \div$ column $b \times 100$.
[3] Column $d \div$ column $a \times 100$.
[4] Includes 7 for whom no occupation was designated.

number of entries above and to the right of the diagonal in table 1 is larger than the number below and to the left. There is movement upward into professionals, managers, and craftsmen from operative and service occupations. Death certificates placed 45 persons in "higher" occupations, socioeconomically, than warranted, and placed only 21 in "lower" occupations.

Frank errors occur throughout table 1 in a somewhat random manner. They are also responsible for the anomaly of persons coded to the correct socioeconomic group but not to the correct occupation within the group, for example, an actor reported as an aviator or a painter reported as a carpenter. While social class is correct for these occupations, environmental exposures may be completely incorrect. Another feature is the variability between groups in proportion of agreement. This is further presented in table 2, which is a summary and extension of part of table 1.

Several comparisons between persons known from occupational history to have specific occupations and those designated with these specific occupations on death certificates are shown in table 2. Columns a, b, and c of the table indicate the order of magnitude of the error in the number of deaths attributed to an occupation compared with what this number should be. It is of some interest to know the numerical error in the numerator of a rate even though a portion of this number may include the wrong persons. The greatest error is for farmers and farm laborers, with over 80 percent excess in the former and 60 percent deficiency in the latter. The obvious explanation would seem to be the upgrading of farm laborers to farmers on death certification. This is partially true but, as can be seen in table 1, misreporting out of and into these categories is not quite that simple and direct. As might be expected, the professional category of occupations has the next highest excess, which is largely a result of upgrading on death certificates, as can be seen in table 1. The other occupational categories are considerably closer to their correct number, with only operatives and service workers deviating as much as 20–25 percent. The actual rates of mortality from death certificates for such categories as managers, clerks, salesper-

sons, and craftsmen will not be severely biased because of excessive or inadequate numbers of deaths attributed to the specific occupations in these categories.

Columns d and e in table 2 show that on the average about 70 percent of those persons designated on death certificates as having a specific occupation will actually have had the occupation. The range is from 45.5 percent for farmers to 81.8 percent for service workers. Column f of table 2 gives the percentage of those persons who have a specific occupation on their occupation history that is identically designated on their death certificates.

The essential proportions for anticipating and interpreting data from the use of occupation on death certificates in the search for abnormal risk of fatal disease are contained in columns c and f of table 2. Column c indicates the degree and direction of bias for mortality rates where, owing to misreporting, the percentage deviation from 100 is unrelated to occupation. For example, occupations in the professional category would be one-third excessive, operatives would be 23 percent deficient, and managers and clerks would be approximately correct. However, in the case of a specific occupation carrying an increased risk of disease, the excess deaths would be retained in the specific occupation on death certificates according to the proportions given in column f.

Conclusion and Summary

At the present it must be admitted that searching for increased risk of specific causes of mortality in specific occupations through death certificate occupational entries is a rather crude and somewhat insensitive method. However, those who are interested have an excellent chance of recognizing increased risks of the order of several fold by this approach. It is also likely that a significant increased risk, when found, will understate the true risk.

Analysis of death rates by cause and by detailed occupation may reveal useful relationships. Cause of death is fairly accurately reported while occupation, not at present being routinely coded or "policed," is less accurate. The bureau of chronic diseases of the California State Department of Health had interviewed a

group of lung cancer patients and controls and had obtained their occupational histories. To study the limitations inherent in the use of death certificates for occupational mortality studies, the occupation reported on the death certificate was later compared with the occupation reported in the interview. The study is based on a chunk sample, but it indicates the kind of results obtainable.

From comprehensive occupational histories, mostly of the lung cancer patients, 304 deceased men aged 35–64 were assigned to the 3-digit occupation codes used by the Bureau of the Census. A similar assignment of occupation codes was made for the occupations reported on the death certificates. Occupation was in agreement when the assigned 3-digit codes were identical. Out of a multiplicity of possible rules for assignment, three were applied to the occupational histories. These rules and the percentage agreements they generated are as follows:

"Usual occupation" method, the National Office of Vital Statistics rule—51.9 percent agreement.

"Last occupation" method, Bureau of the Census rule—69.7 percent agreement.

"Five-year exposure occupation" method, our own rule—69.4 percent agreement.

Despite a phrase on the death certificate defining usual occupation, the death certificate entries best represent last occupation. Because of this situation, because death rates must be computed by using death certificate populations for the numerator and census populations for the denominator, and because the denominator occupations are defined by the census rule for last occupation, we used the last occupation rule for the balance of the study.

By grouping occupations into 10 broad groups, though retaining 3-digit agreement, we obtained some idea of the variability of misreporting. Since it cannot be assumed that this variability is random, the specific net misreporting for each group must be used with caution. By considering the effects of positive association between a cause of death and an occupation, it seems evident that strong positive associations will show themselves despite considerable misreporting.

REFERENCES

(1) Pott, P.: Chirurgical observations relative to the cataract, the polypus of the nose, the cancer of the scrotum, the different kinds of ruptures, and the mortification of the toes and feet. London, Hower, Clarke & Pollins, printed by T. J. Carnegie, 1775, p. 63.

(2) Machle, W., and Gregorius, F.: Cancer of respiratory system in the United States chromate producing industry. Pub. Health Rep. 63: 1114–1127. Aug. 27, 1948.

(3) Morris, J. N., Heady, J. A., Raffle, P. A. B., Roberts, C. G., and Parks, J. W.: Coronary heart disease and physical activity of work. Lancet 267: 1053–1057, 1111–1120, Nov. 21 and 28, 1953.

(4) Breslow, L., Rasmussen, G., Hoaglin, L., and Abrams, H. K.: Occupation and cigarette smoking as factors in lung cancer. Am. J. Pub. Health 44: 171–181, February 1954.

(5) United States Bureau of the Census: Census of population, 1950. A report of the seventeenth decennial census of the United States. Characteristics of the population. Pt. 9. District of Columbia. Washington, D. C., U. S. Government Printing Office, 1952.

(6) Strauss, R. C.: Guide for reporting occupation and industry on death certificates. Washington, D. C., U. S. Government Printing Office, 1949, p. 4.

(7) United States Bureau of the Census: Alphabetical index of occupations and industries. Washington, D. C., U. S. Government Printing Office, 1950.