

How good are studies based on mortality statistics? This investigation was undertaken to determine the quality of diagnostic evidence supporting the accuracy of cause-of-death statements.

INQUIRY INTO DIAGNOSTIC EVIDENCE SUPPORTING MEDICAL CERTIFICATIONS OF DEATH

Iwao M. Moriyama, Ph.D., F.A.P.H.A.; William S. Baum, M.D.; William M. Haenszel, F.A.P.H.A.; and Berwyn F. Mattison, M.D., F.A.P.H.A.

STATISTICS ON causes of death have been used extensively in the epidemiological study of diseases. Because of the systematic coverage of deaths in the total population, mortality statistics have provided useful information for epidemiological studies. However, the value of such studies has been questioned from the standpoint of inaccuracies in the statements of causes of death. The studies on the accuracy of cause-of-death statements have taken clinical or pathological reports as the point of departure for comparison with the medical returns on the death certificate.* 1-3

It is difficult to generalize from the results of these studies in the interpretation of general mortality statistics. By necessity, these studies have been limited to deaths on which hospital records or autopsy reports were available. Since less than half of all deaths occur in hospitals and only about 15 per cent of all deaths go to autopsy, studies based on such data may be subject to considerable bias. Hospital deaths tend to be weighted with surgical cases, acute illnesses, and illnesses more difficult to diagnose. Pathological findings are sought more for

deaths where the clinical diagnosis is not clear, or where diagnostic confirmation is desired. Thus, the selection of hospital deaths, and more particularly deaths where autopsies are performed, should lead to greater disagreement than would appear in a hypothetical situation where cause-of-death statements for all deaths are compared with clinical or pathological findings. Since, in practice, clinical and pathological reports are not available for all deaths, one alternative is to start with the death certificate and query the physician. In this manner, his knowledge and subjective impressions can be utilized for the interpretation of cause-of-death statements.

Material and Methods

To explore the use of death certificates in delineating possible relations of smoking, residence, and work histories to cancer of the lung,⁵ a sample of 1,837 deaths occurring in Pennsylvania during the three months of 1956 was studied. As part of this pilot study, additional information was collected from certifying physicians to determine the possibilities of developing measures of quality of medical certifications and to ascertain the problems involved in the conduct of such a study. These investigations were conducted in cooperation with

* One exception to this approach is the study by Nicoll and Bellows,⁴ who took a sample of death certificates and queried the certifying physician about the presence of syphilis, tuberculosis, alcoholism.

the Pennsylvania State Department of Health.

The death certificates selected for this purpose came from the 10 per cent Current Mortality Sample (CMS) sent to the National Office of Vital Statistics, a representative sample of all deaths in the state. The causes covered in this investigation included tuberculosis, malignant neoplasms, diabetes, all cardiovascular-renal diseases, influenza and pneumonia, and cirrhosis of the liver. Such causes as accidents, suicides, and homicides were excluded, and the CMS sample augmented by inclusion of all deaths from cancer of the respiratory system occurring during the three-month period. For cardiovascular-renal causes, not all certificates in the CMS sample were queried; depending on age and race, one-fourth to one-half of the certificates of death for those over age 40 were selected for study. The tabular results for the total sample and certain subgroups of causes have been adjusted to take account of the differential rate of sampling. The adjusted totals represent one-tenth of the deaths from the causes selected for study in Pennsylvania in the three-month period.

Questionnaires were sent to physicians signing the death certificates asking for information on the following points: (a) diagnostic methods, with pertinent findings, on which the medical certification of death was based; (b) an expression of his certainty of the diagnosis entered on the medical certification; (c) a revised medical certification if his opinion on diagnosis had changed since signing the death certificate. When the certifier indicated that his source of information was another physician who had previously attended the patient, a query was sent to that person as well. Follow-up letters were sent as indicated. Completed returns were received from physicians for 96 per cent of the deaths queried. Deaths certified by coroners and medical examiners were followed in

the manner described above, except that expressions of certainty of diagnosis were not requested.

The returns were reviewed along with the original cause-of-death statement under the direction of an internist (WSB) and rated on the following points: (a) quality (type and amount) of supporting diagnostic information; (b) consistency of medical certification with the diagnostic evidence; and (c) the physician's opinion of certitude of diagnosis. The reviewing internist also entered his impression of certainty of the diagnosis. The latter rating was not independent of that stated by the certifier. In many instances, the physician's opinion was taken into account with the other diagnostic evidence in assigning a rating. The ratings were later checked by listing individual deaths in cause order with respect to specific diagnostic methods, quality of supporting diagnostic information, consistency of medical certification and certitude of diagnosis. The combined information was then reviewed and checked for internal consistency.

Results

Medicolegal Certifications—Table 1 shows, by cause, the number and proportion of medicolegal certifications. Except for cardiovascular-renal diseases, the number of deaths certified by the medical examiner and coroners for the causes studied was insignificant. The problems of medicolegal certifications cannot be regarded as an important contributing source of error in mortality tabulations for nonviolent deaths other than those from cardiovascular-renal diseases.

Over 90 per cent of the deaths from natural causes certified by the medical examiner and coroners in the three-month period were ascribed to diseases of the cardiovascular-renal system. The number of deaths signed out by the

Table 1—Death Certificates with Cause of Death Certified by Coroner or Medical Examiner: Pennsylvania Sample, Three-Month Period, 1956

(Numbers after causes of death are those of the Sixth Revision of the International Lists, 1948)

Cause of Death	Total Death Certificates	Cause of Death Certified by Coroner or Medical Examiner	
		Number	Per cent of Total
Total Weighted Sample *	2,122	232	10.9
Tuberculosis, all forms	27	2	7.4
Malignant neoplasms, including neoplasms of lymphatic and hematopoietic tissues *	443	10	2.3
Diabetes mellitus	61	1	1.6
Major cardiovascular-renal diseases * 330-334, 400-456, 590-593	1,406	212	15.1
Vascular lesions affecting central nervous system *	254	18	7.1
Rheumatic fever and rheumatic heart disease *	41	3	7.3
Arteriosclerotic heart disease so described....	297	37	12.5
Heart disease specified as involving coronary arteries *	395	124	31.4
Other heart disease except hypertension * 421, 422, 430-434	167	7	4.2
Hypertensive diseases *	150	22	14.7
Other cardiovascular and renal diseases * 450-456, 590-593	102	1	1.0
Influenza and pneumonia, except pneumonia of newborn	31	1	3.2
All other	154	6	3.9

* Figures adjusted to represent one-tenth of the deaths for the three-month period, that is, the number of deaths in the Current Mortality Sample.

medical examiner and coroners as cardiovascular-renal diseases represented 15 per cent of the total recorded for this cause in Pennsylvania. A surprising feature is that medicolegal certifications account for only 7 per cent of deaths from vascular lesions of the central nervous system, since this cause often results in sudden death. Another cause frequently linked with sudden death, coronary heart disease, does exhibit a high proportion of medicolegal certifications, 31 per cent. The latter figure is consistent with experience recorded elsewhere. This apparent anomaly arises because persons who suffer cerebrovascu-

lar accidents often have symptoms placing them under more or less continuous medical supervision. Coronary disease, on the other hand, may occur among persons with no recent history of medical attention. These data clearly indicate that the care with which the medical examiner and coroner discharge their legal responsibilities can significantly affect the statistics for coronary and hypertensive heart diseases.

Quality of Diagnostic Information—The criteria used for rating the quality of diagnostic information differed for specific diseases and are summarized in Table 2. The results of these ratings

Table 2—Classification of Quality of Supporting Diagnostic Information

Cause of Death (Int. List Nos.)	Very Good	Good	Sketchy
Tuberculosis (001-019)	Autopsy or operation and biopsy or x-ray with positive sputum or culture	X-ray with increasing area of destruction	Clinical impression; suggestive x-ray findings
Malignant neoplasms (140-205)	Autopsy or operation or endoscopy with microscopic confirmation*	Positive cytology alone or in combination with other nonmicroscopic findings or operation or definite x-ray findings, without microscopic confirmation or biopsy of a metastatic site or laboratory findings for specific sites	Clinical impression; suggestive x-ray findings
Diabetes mellitus (260)	Diabetic complications or typical clinical findings supported by laboratory evidence	Strongly suggestive clinical and laboratory findings	No detailed support of clinical diagnosis other than reference to urinalysis
Vascular lesions affecting central nervous system (330-334)	Autopsy or typical clinical findings or good clinical history supported by lumbar puncture or history of previous attacks	Strongly suggestive clinical findings not supported by lumbar puncture or history of previous attack	No detailed support of clinical diagnosis
Rheumatic heart disease (410-416)	Autopsy or combination of at least two of the following items; history of rheumatic fever, clinical picture (heart murmur), definite EKG findings	Typical history of rheumatic fever or strongly suggestive clinical or EKG findings	No detailed support of clinical diagnosis
Other forms of heart disease (420-447)	Autopsy or clinical findings supported by definite EKG findings or for hypertensive disease by definite x-ray findings	Typical clinical signs alone or combination of suggestive clinical and EKG or x-ray findings	No detailed support of clinical diagnosis
Influenza and pneumonia (480-493)	Autopsy †	Clinical findings supported by x-ray or culture findings	No detailed support of clinical diagnosis

* Removal of an organ was interpreted as implying microscopic confirmation.

† While for lobar pneumonia a combination of positive clinical and culture findings would have been classified as "very good," no such cases were encountered in the sample studied.

by cause of death are presented in Table 3. Since the major concern is with the interpretation of general mor-

tality data, all analyses are presented in terms of the underlying cause as stated on the original death certificate.

• **Table 3—Deaths from Specified Causes by Quality (Type and Amount) of Supporting Diagnostic Information: Pennsylvania Sample, Three-Month Period, 1956**
(Numbers after causes of death are category numbers of the Sixth Revision of the International Lists, 1948)

Cause of Death	Total Death Certificates	Per cent Distribution				
		Total	Very Good	Good	Sketchy	No Report
Total Weighted Sample *.....	2,122	100.0	39.1	18.5	38.5	3.9
Tuberculosis, all forms.....001-019	27	100.0	74.1	0	14.8	11.1
Malignant neoplasms, including neoplasms of lymphatic and hematopoietic tissues *.....140-205	443	100.0	68.4	12.9	15.3	3.4
Malignant neoplasm of digestive organs and peritoneum150-159	191	100.0	65.4	12.0	18.3	4.3
Of stomach151	49	100.0	51.0	20.4	24.5	4.1
Of large intestine including rectum.....153, 154	97	100.0	72.2	9.3	13.4	5.2
Of pancreas157	16	100.0	56.3	12.5	31.3	0
Malignant neoplasm of respiratory system *.....160-165	59	100.0	62.7	18.6	15.3	3.4
Of trachea, and of bronchus and lung specified as primary162	21	100.0	66.7	19.0	9.5	4.8
Of lung and bronchus, unspecified as to whether primary or secondary163	33	100.0	57.6	21.2	18.2	3.0
Malignant neoplasm of breast.....170	32	100.0	68.8	15.6	12.5	3.1
Malignant neoplasm of uterus.....171-174	28	100.0	85.7	0	14.3	0
Malignant neoplasm of male genital organs....177-179	21	100.0	61.9	14.3	19.0	4.8
Malignant neoplasm of urinary organs.....180, 181	16	100.0	87.5	6.3	6.3	0
Malignant neoplasm of lymphatic and hematopoietic tissues200-205	28	100.0	89.3	10.7	0	0
Diabetes mellitus260	61	100.0	42.6	24.6	31.1	1.7
Major cardiovascular-renal diseases * 330-334, 400-456, 590-593	1,406	100.0	28.7	19.8	47.4	4.1
Vascular lesions affecting central nervous system *330-334	254	100.0	20.9	23.2	51.2	4.7
Rheumatic fever and rheumatic heart disease *400-402, 410-416	41	100.0	58.5	19.5	14.6	7.4
Arteriosclerotic heart disease, including coronary disease420.0, 420.1	692	100.0	29.8	17.6	49.0	3.6
Other heart disease except hypertension *421, 422, 430-434	167	100.0	17.4	16.2	64.1	2.3
Hypertensive diseases *440-447	150	100.0	30.7	27.3	36.0	6.0
Influenza and pneumonia, except pneumonia of newborn480-493	31	100.0	16.1	25.8	51.6	6.5

* Figures adjusted to represent one-tenth of the deaths for the three-month period, that is, the number of deaths in the Current Mortality Sample.

The kind and amount of diagnostic information reported as being available to the medical certifier varied considerably. In 4 per cent of the deaths, no information was volunteered at all. In 38 per cent, the diagnostic data given were sketchy. In the remainder, 58 per cent, the information was considered very good (39 per cent) or good (19 per cent). In interpreting these results, it should be pointed out that the proportion of "very good" and "good" ratings are minimum figures since the physician may have had more information at his disposal which was not communicated

on the query form. Also, the computed percentages do not exclude deaths for which no information was reported.

The picture for total deaths does not differ greatly from that for physicians' certifications alone. This is not surprising since physicians' certifications account for the bulk of total deaths, 91 per cent. It may be more surprising that for the important cardiovascular-renal disease category the proportion of "very good" medicolegal certifications, 20 per cent, approached that for physicians' certifications, 30 per cent. The apparent reason, in Pennsylvania at least,

is that the coroner or medical examiner often consults the attending physician, if there is one, before making out the death certificate; or if the reasons for death seem unclear, he orders an autopsy performed.

The quality of reported information on diagnostic methods varies by cause. It was better for malignant neoplasms than for other diseases. In 68 per cent of the deaths attributed to cancer, the diagnostic information was "very good," meaning that it was based on micro-

scopic examination of tissues. This result is consistent with evidence provided by cancer morbidity surveys in 10 metropolitan areas⁶ and in Iowa.⁷ While the numbers studied are not large enough to yield more than suggestive results, the diagnostic evidence for malignant neoplasms appears weakest for stomach and pancreas, a finding in agreement with prevailing clinical opinion. The diagnoses of malignant neoplasms of the lymphatic and hematopoietic tissues were particularly well supported, being rated

Table 4—Consistency of Medical Certification with Diagnostic Evidence by Cause of Death: Pennsylvania Sample, Three-Month Period, 1956

(For category numbers of causes of death see Table 3)

Cause of Death	Total Death Certificates	Per cent Distribution				
		Total	Most Probable Diagnosis	Another Diagnosis Equally Probable	Another Diagnosis Preferred	No Diagnostic Information
Total Weighted Sample *	2,122	100.0	78.7	12.7	4.7	4.9
Tuberculosis, all forms	27	100.0	85.2	0	3.7	11.1
Malignant neoplasms, including neoplasms of lymphatic and hematopoietic tissues *	443	100.0	85.8	5.0	5.6	3.6
Malignant neoplasm of digestive organs and peritoneum	191	100.0	82.2	7.9	5.8	4.1
Of stomach and large intestine except rectum	120	100.0	83.3	5.0	6.7	5.0
Of rectum	26	100.0	88.5	3.8	3.8	3.9
Of biliary passages and of liver (stated to be primary site)	13	100.0	92.3	0	0	7.7
Of pancreas	16	100.0	68.8	31.3	0	0
Malignant neoplasm of respiratory system *	59	100.0	88.1	5.1	1.7	5.1
Of larynx	4	100.0	100.0	0	0	0
Of trachea and of bronchus and lung specified as primary	21	100.0	90.5	4.8	0	4.7
Of lung and bronchus, unspecified as to whether primary or secondary	33	100.0	84.8	6.1	3.0	6.1
Malignant neoplasm of breast and female genital organs	68	100.0	94.1	1.5	1.5	2.9
Of breast	32	100.0	93.8	0	3.1	3.1
Of uterus	28	100.0	96.4	3.6	0	0
Malignant neoplasm of male genital organs	21	100.0	90.5	4.8	0	4.7
Malignant neoplasm of urinary organs	16	100.0	100.0	0	0	0
Malignant neoplasm of lymphatic and hematopoietic tissues	28	100.0	100.0	0	0	0
Leukemia and aleukemia	13	100.0	100.0	0	0	0
Lymphomas and myelomas	15	100.0	100.0	0	0	0
Diabetes mellitus	61	100.0	86.9	8.2	3.3	1.6
Major cardiovascular-renal diseases *	1,406	100.0	78.2	14.4	3.3	4.1
Vascular lesions affecting central nervous-system *	254	100.0	85.0	8.7	1.6	4.7
Rheumatic fever and rheumatic heart disease *	41	100.0	82.9	9.8	0	7.3
Arteriosclerotic heart disease so described and heart disease specified as involving coronary arteries *	692	100.0	78.6	16.5	1.3	3.6
Other heart disease except hypertension *	167	100.0	66.5	19.8	11.4	2.3
Hypertensive diseases *	150	100.0	76.0	16.0	2.0	6.0
Other cardiovascular-renal diseases *	102	100.0	78.4	5.9	11.8	3.9
Influenza and pneumonia, except pneumonia of newborn	31	100.0	64.5	19.4	9.7	6.4
All other	154	100.0	61.7	21.4	14.3	2.6

* Figures adjusted to represent one-tenth of the deaths for the three-month period, that is, the number of deaths in the Current Mortality Sample.

as very good or good for every death from the leukemias and lymphomas.

Consistency of Medical Certification—The medical certifications were then rated with respect to consistency with the diagnostic information reported. The results of this rating are shown in Table 4. This review indicated that the reported causes were the most probable diagnoses in 79 per cent of the deaths, other diagnoses appeared equally probable in 13 per cent, and in 5 per cent another diagnosis was preferred. In about 4 per cent of the cases, no diagnostic evidence was presented. The certifications appeared more open to question in the medicolegal cases where 61 per cent of the 232 certifications were categorized as the most probable diagnosis. Virtually all of the 116 deaths for which another diagnosis was preferred resulted from changes in the certifier's diagnosis, usually attributable to additional information obtained after the death certificate was filed. The certifying physician changed his cause-of-death statement for 5 per cent of the deaths in the sample. This proportion did not vary significantly by cause. The reviewer agreed with the certification change in most instances.

Diabetes, malignant neoplasms, and tuberculosis ranked high in representing the most probable diagnosis according to the medical review of the diagnostic evidence. Over 85 per cent of deaths attributed to these causes were judged to be the most probable diagnosis.

There seemed to be little question about accepting the diagnosis of malignant neoplasm for most sites. Stomach and pancreas were two exceptions, where the evidence presented was much less convincing in the support of the diagnosis.

While the quality of diagnostic information for the cardiovascular-renal deaths leaves something to be desired, it is clear from Table 4 that the medical certifications given represent the most

probable diagnoses, and, as a rule, are preferable to other diagnoses which might be advanced. The figures may understate the quality of medical certifications for the total of the cardiovascular-renal diseases since the judgments were made in terms of the specific diagnoses reported. Some classified as "another diagnosis equally probable" or "another diagnosis preferred" could well have been the "most probable diagnosis" had the cause-of-death statement referred to another related disease within the cardiovascular-renal system.

Quality of Diagnostic Inference—As previously stated, the reviewer scored (a) the amount and kind of diagnostic information reported, and (b) the reasonableness of the inference drawn by the medical certifier on the basis of the diagnostic data given. In addition, the reviewer gave his own opinion as to the certitude of the diagnosis reported on the basis of the information at hand. The three scores were then combined and grouped into four categories, namely, solidly established diagnosis, reasonable diagnosis, diagnosis in doubt, and diagnosis probably wrong. These categories resemble the measures presented in Table 4. However, because the combined index also took into account the amount and kind of diagnostic data reported and the reviewer's opinion of certitude, it may be considered a more balanced assessment.

Table 5, which excludes medicolegal cases, shows the distribution of the reviewer's evaluation by cause. For all causes in the Pennsylvania sample, 43 per cent of the medical certifications appear to have been solidly established, and an additional 36 per cent seemed to be reasonable diagnoses. In 10 per cent there was some question as to whether the cause of death statement was right or wrong, and for 8 per cent of the death certificates, the cause, as stated originally, was probably wrong. In 4 per cent no information on diagnos-

Table 5—Reviewer's Evaluation of Diagnostic Information on the Deaths Certified by Physicians: Pennsylvania Sample, Three-Month Period, 1956

(Excludes 232 medicolegal cases. Deaths for which no information was reported are included in totals. For category numbers of causes, see Table 3)

Cause of Death	Total Number	Per cent Distribution				
		Total	Diagnosis of Cause of Death*			
			Solidly Established	Reasonable	In Doubt	Probably Wrong
Total Weighted Sample †	1,890	100.0	42.7	36.0	9.6	7.5
Tuberculosis, all forms	25	100.0	72.0	12.0	0	4.0
Malignant neoplasms, including neoplasms of lymphatic and hematopoietic tissues †	433	100.0	67.7	14.1	7.4	7.2
Malignant neoplasm of digestive organs and peritoneum of stomach	185	100.0	65.4	13.0	9.2	8.1
Malignant neoplasm of respiratory system †	47	100.0	44.7	21.3	10.6	19.1
Of trachea, and of bronchus and lung specified	58	100.0	61.1	21.8	12.4	5.0
as primary	21	100.0	61.9	23.8	4.8	4.8
Of lung and bronchus, unspecified as to whether primary or secondary	33	100.0	57.6	21.2	9.1	6.1
Malignant neoplasm of breast and female genital organs	66	100.0	81.8	9.1	4.5	1.5
Malignant neoplasm of male genital organs	20	100.0	65.0	25.0	5.0	0
Malignant neoplasm of urinary organs	16	100.0	87.5	6.3	6.3	0
Malignant neoplasm of lymphatic and hematopoietic tissues	28	100.0	89.3	10.7	0	0
Diabetes mellitus	60	100.0	50.0	36.7	6.7	5.0
Major cardiovascular-renal diseases †	1,194	100.0	32.8	46.5	10.3	6.0
Vascular lesions affecting central nervous system †	236	100.0	31.8	53.8	7.2	2.1
Rheumatic fever and rheumatic heart disease	38	100.0	58.0	23.7	7.9	2.6
Arteriosclerotic heart disease so described †	260	100.0	30.0	51.9	9.2	3.5
Heart disease specified as involving coronary arteries †	271	100.0	39.1	45.0	10.7	3.0
Hypertensive diseases †	128	100.0	32.8	40.6	16.4	3.1
Other cardiovascular-renal diseases †	261	100.0	26.4	42.1	11.1	17.2
Influenza and pneumonia, except pneumonia of newborn	30	100.0	20.0	40.0	16.7	16.7
All other	148	100.0	45.9	18.9	12.2	20.3

* See footnote on Table 6 for criteria.

† Figures adjusted to represent one-tenth of the deaths for the three-month period, i.e., the number of deaths in the Current Mortality Sample.

tic procedures and methods was given to permit an evaluation. As stated previously, most of the decisions of "probably wrong" diagnoses were concurred in by the reporting physician.

For malignant neoplasms of all sites 65 per cent may be regarded as solidly established. This is slightly lower than the proportion of deaths with "very good" supporting diagnostic information since the "very good" data may have proved indecisive. An additional 14 per cent fell into the category of "reasonable diagnosis."

In contrast, the cardiovascular-renal disease had a much lower proportion of solidly established diagnoses, 33 per cent, compensated by a higher pro-

portion of reasonable diagnoses, 47 per cent. The results indicated more uncertainty concerning the accuracy of diagnoses in the hypertensive diseases and the residual group of cardiovascular-renal diseases. The query form was probably weakest in eliciting clinical evidence which, in part, may account for the lower proportion of solidly established and reasonable diagnoses for these rubrics.

Table 6 compares for all causes the reviewer's evaluation and the physician's certitude of diagnosis given in response to a specific question on the query form. This comparison should not be construed as representing two independent ratings, since, in some instances,

Table 6—Reviewer's Evaluation as Compared with Physician's Opinion of Certitude of Diagnosis: Pennsylvania Sample, Three-Month Period, 1956

Reviewer's Evaluation	Total Hours	Physician's Opinion of Certitude			
		Positive	Reasonably Certain	Somewhat Uncertain	Very Uncertain
Total	1,789	1,032	580	28	13
Solidly established	775	632	119	1	6
Reasonable diagnosis	664	283	344	15	3
In doubt	168	64	75	11	1
Probably wrong	106	53	42	0	3

Ratings based upon following 3-digit combinations of factors:

Solidly established 111, 211, 112

Reasonable 212, 312, 311

In doubt 113, 213, 313, 323, 123, 122, 132

Probably wrong 324, 333, 334, 134, 133, 234, 124, 224, 233, 223

Where the first digit of the 3-digit combination denoted the diagnostic information available to certifier

1. Very good

2. Good

3. Sketchy

the second digit of the 3-digit combinations denoted the validity of diagnosis as stated on death certificate

1. Most reasonable inference

2. Another diagnosis equally probable

3. Another diagnosis preferred

and the third digit of the 3-digit combinations denoted the certitude of reviewer of the diagnosis

1. Positive

2. Reasonably certain

3. Somewhat uncertain

4. Uncertain

the reviewer took account of the physician's opinion in his evaluation. The cross-tabulation was restricted to cases for which the medical certification remained unchanged, since in the other situation it was uncertain whether the physician's certitude related to the original or revised certification. Some anomalies such as the reviewer's evaluation of "solidly established" or "reasonable diagnosis" in the presence of a physician's opinion of "somewhat uncertain" or "uncertain" can be usually accounted for by data available to the reviewer and not to the certifier. For example, the certifier might have had at his disposal only hearsay evidence for a patient referred to him for terminal care. In writing to the original diagnostic sources for further information, unequivocal evidence was elicited in many instances.

A comparison of the reviewer's evaluation and the certifying physician's

opinion of certitude of his diagnoses shows some lack of agreement. In general, the certifier was more certain of the accuracy of the diagnoses than the reviewer. For example, in 90 per cent of the medical certifications the physician was positive or reasonably certain of the diagnoses, while the reviewer rated 80 per cent of the diagnoses as "solidly established" or "reasonable." One likely explanation deals with a common problem in the recording of information. The certifier may have had in mind unreported information which played an important part in firming up his opinion of certitude. Still another possibility is that the reviewer did not receive from the condensed report quite the impression which the certifier meant to convey. These factors may account for some of the 95 deaths, or about 5 per cent of the physician's certifications where the reviewer felt that the medical certification was probably incorrect de-

spite the certifier's expression of confidence in his diagnosis. As mentioned before, some of those judged probably incorrect could have been correct had the cause-of-death statement referred to another disease within the same organ system. Despite the discrepancies noted, the table still gives a strong impression of basic agreement between the certifier and reviewer.

Other Factors Associated with Diagnostic Quality—The reviewer's evaluation of diagnostic quality were then examined with respect to several characteristics of the decedent and the certifier, including the decedent's sex, age, and place of residence, and the certifier's age, specialty, place of practice, and length of attendance on the case.

Further analysis of diagnostic quality shows that the study of cardiovascular-renal mortality differentials by sex, for example, should take into account the better quality of diagnoses for males. The proportion of diagnoses solidly established is 36 per cent for males and 30 per cent for female decedents. In the more usual situation where the interest is in reasonable diagnosis or better, this factor does not appear to be important. For other causes, the sex difference in the proportion of solidly established diagnoses was not important.

By age, the quality of diagnosis falls off for deaths in the older ages. After age 65, the quality becomes progressively poorer for malignant neoplasms. This suggests that the analysis of cancer mortality might emphasize data under 75 years of age. For the cardiovascular-renal diseases, there is not the same degree of change by age. However, there is a sharp dip for the age group 45-54 years occurring in both sexes, the reasons for which may merit further study. The poorer quality of diagnostic data in the older ages is probably linked with the medical problems presented by these patients. The patient's condition may be such that ag-

gressive attempts to reach a definitive diagnosis are not made because they would not be helpful in treatment.

Data by place of residence show slight differences by population size, but these do not seem large enough to account for much of the observed urban-rural differences in mortality.

The results by length of time the certifier had attended the patient before death were inconclusive, partly because of the proportion on which this item was not answered. One might have expected a strong association between length of attendance and quality of diagnosis. However, the present study made it clear that in many instances the physician last in attendance was able to take account of diagnostic information developed by others.

There seemed also to be a marked urban-rural difference in the quality of diagnoses according to the physician's place of practice. About 80 per cent of the diagnoses of urban medical practitioners were rated as "solidly established" or "reasonable" while 70 per cent of the diagnoses of their rural colleagues fell into that category. The same order of difference was apparent for diagnoses involving malignant neoplasms and cardiovascular-renal diseases.

Discussion

Within the limits of prevailing medical knowledge, the accuracy and completeness of cause of death statements depend upon (1) availability of pertinent diagnostic information, (2) diagnostic acumen on the part of the physician, and (3) the manner in which diagnoses are reported on the death certificate. This paper deals with the first factor; the third will be discussed in another paper.

The problem of diagnostic acumen, although a basic factor, was not evaluated in this study. In general, the reported diagnoses were accepted if suffi-

ciently supported by proper diagnostic tests and methods. Errors are possible even in those cases where the diagnoses are rated as "solidly established."

Other investigators have approached the diagnostic accuracy problem by comparing the medical certification with findings for deaths coming to autopsy. However, it is frequently difficult to relate the pathological findings to each other or to the sequence of events leading to death. Furthermore, diagnoses of clinical entities cannot always be established by autopsy findings. For example, James, et al.,³ reported that the use of autopsy data, including a clinical summary, as a standard of reference decreased the proportion of deaths assigned to such causes as diabetes and hypertensive heart diseases and increased the proportion attributable to arteriosclerosis. In discussion, they suggested that "the lack of clinical information or the lesser attention given to this group of diseases by pathologists may be a partial answer for these differences. . . ." They commented further that "in certifying these deaths, the physician may have given more weight to clinical information than did the pathologist." Such a course of action is understandable and appears proper. It would seem highly undesirable to emphasize the pathological aspects of death without giving due weight to the clinical findings.

In practice, the comparison of death certificates with autopsy findings has suffered from a defect similar to that encountered in this study—that is, the third-party review takes account only of recorded information. Unless the reported data are complete and accurate, serious problems of misinterpretation will arise.

In the present study, the internist took into consideration both the reported clinical and pathological information. This approach has its shortcomings, some of which might be overcome by

modifying the data collection method. One improvement would be to follow the pattern of clinical or clinical-pathological conferences where the principal disciplines associated with the case would present the facts and discuss the significance of the findings. This method is feasible only for deaths occurring within a single institution. To permit generalizations of findings applicable to all deaths within a study area, a broader coverage of deaths, not restricted to a single institution is needed. Here the query approach for a sample of deaths could be supplemented by interviews conducted by an internist with persons who have some knowledge of medical aspects of the case.

Although only one internist conducted the review in this study, provision for additional reviewers would probably not alter the results substantially. More important is the choice of criteria for the evaluation of the reported diagnostic methods and findings. The criteria adopted for this study appeared sufficiently discriminating and did not give rise to many classification difficulties.

While tentative and preliminary, the results of this approach appear promising enough to warrant further trial and development. The method holds special promise for the interpretation of time trends and regional differences in mortality for specific causes of death. When confronted with mortality differentials, the question arises as to whether the observed difference is real or a reporting artifact. In this connection, one of the more difficult problems has been in assessing the reliability of diagnostic information on death certificates. The systematic collection of additional data of the type reported here should be of material assistance in indicating whether diagnostic differences could be an important complicating factor in the interpretation of the mortality data.

The results reported for Pennsylvania may turn out to be descriptive for other

northern states. More data are needed and further work should be done on a national scale to explore the nature of regional differences. Some thought should be given to the periodic collection of data for the interpretation of national trends in mortality for specific causes. The results suggest that the cardiovascular-renal diseases present the greatest problems in interpretation and special attention should be given to investigations in that area. The effectiveness of a supplementary query program can unquestionably be improved by restricting the inquiry to a single disease.

This approach to the collection of supplementary data offers some additional advantages. For example, the Pennsylvania experience indicates definitely that it is feasible to collect data on histologic type of cancer of specific sites as part of a supplementary query. For reasons which will not be elaborated here, it cannot be expected that such information can be obtained from the medical certification.

Absolute and unflinching accuracy of ante-mortem diagnoses are unattainable goals. Although there are opportunities to secure additional information post mortem, insistence on absolute precision

of diagnosis for the purposes of preparing cause-of-death statements appears wholly unreasonable. There should be a more general recognition that mortality data are not precise measures but that they are useful in suggesting leads to be elaborated by other study approaches. The results of this study suggest that the diagnostic data for many disease categories in Pennsylvania, at least, are more than adequate for this purpose.

REFERENCES

1. Swartout, H. O., and Webster, R. J. To What Degree Are Mortality Statistics Dependable? *A.J.P.H.* 30:811 (July), 1940.
2. Pohlen, K., and Emerson, H. Errors in Clinical Statements of Causes of Death. *Ibid.* 32:251 (Mar.), 1942; 33:505 (May), 1943.
3. James, G.; Patton, R. E.; and Heslin, A. S. Accuracy of Cause-of-Death Statements on Death Certificates. *Pub. Health Rep.* 70:39 (Jan.), 1955.
4. Nicoll, M. Jr., and Bellows, M. T. Effect of a Confidential Inquiry on the Recorded Mortality from Syphilis and Alcoholism. *A.J.P.H.* 24:813 (Aug.), 1934.
5. Haenszel, W.; Moriyama, I. M.; and Sirken, M. G. "A Proposed Study for Extending the Scope and Improving the Quality of Mortality Data," in *Improving the Quality of Statistical Surveys.* p. 57. Washington, D. C.: American Statistical Association, 1956.
6. Dorn, H. F., and Horn, J. I. The Reliability of Certificates of Death from Cancer. *Am. J. Hyg.* 34:12 (July), 1941.
7. Haenszel, W.; Marcus, S. C.; and Zimmerer, E. G. Cancer Morbidity in Urban and Rural Iowa. *Pub. Health Monograph No. 37, PHS Pub. No. 462.* 1956.

Dr. Moriyama is chief, Mortality Analysis, National Office of Vital Statistics; Dr. Baum, formerly assistant director, National Cancer Institute, is now with the Indian Health Service, Public Health Service, Washington, D. C.; Mr. Haenszel is head, Demography Section, Biometry Branch, National Cancer Institute, Bethesda, Md.; and Dr. Mattison, formerly secretary of health, Pennsylvania State Health Department, is executive secretary, American Public Health Association.

This paper was presented before a Joint Session of the Epidemiology and Statistics Sections of the American Public Health Association at the Eighty-Fifth Annual Meeting in Cleveland, Ohio, November 11, 1957.