Teitelbaum
1-20-89

PLAINTIFF'S EXHIBIT 1206
PENGAD-Bayonne, N. J.

Death Certificate - Problems

Incidence, Prevalence, and Evidence

Scientific Problems in Epidemiologic Statistics for the **Occurrence** of *Cancer*

ALVAN R. FEINSTEIN, M.D.
New Haven, Connecticut
and
West Haven, Connecticut
JOHN M. ESDAILE, M.D.*
New Haven, Connecticut

Major changes in policy for health, nutrition, and industrial safety have been proposed because of the "epidemic of apprehension" produced by statistics for the occurrence rates of cancer. The statistical information, however, contains gross violations of epidemiologic principles and scientific standards for credible evidence. The calculations often depend on an improper linkage of numerators and denominators; and the calculated rates, assembled from reports of overtly detected cases of cancer, represent neither incidence nor prevalence. Many of the secular or regional changes in rates are due to corresponding changes in the availability and dissemination of diagnostic technology, but the technologic changes have not been adequately evaluated or investigated. Improvements will require drastically altered approaches to the use of death certificates, tumor registries, and clues from necropsy data.

Hardly a day goes by without another frightening announcement about the occurrence rates of cancer. They are regularly reported as increasing in groups of people demarcated according to age, gender, race, or residence in certain states, cities, or other geographic regions. The demographic demarcations are often accompanied by statistics showing that the increased rates are associated with diverse features of environment or nutrition, various types of chemicals, certain occupations and industries, or other suspected etiologic culprits. Receiving these frequent announcements, the public has developed what Lewis Thomas [1] has called an "epidemic of apprehension." Cancer appears to lurk in every aspect of daily life: the air we breathe, the water we drink, the food we eat, the homes we live in, the substances we touch, and the work we do.

The purpose of this essay is to examine the methods with which the cancer statistics are assembled, to note the many problems that impair the scientific credibility of the data, and to suggest methods of improving the defects.

DEFINITIONS OF PREVALENCE AND INCIDENCE

Like other forms of epidemiologic data, statistics for the occurrence of cancer are usually reported as rates of *incidence* or *prevalence*. These two terms are constantly used to express distinctions in the many occurrence rates for which epidemiology is famous; and a knowledge of the distinctions is a professional hallmark of epidemiologists. *Prevalence* is defined as the relative frequency with which a particular event is noted to be present, at a single point in time, in a particular population or group of people. *Incidence* is defined as the relative frequency with which the event *newly* appears as the persons at risk are followed forward during a subsequent interval of time. For example, if we examine a general

From the Clinical Epidemiology Unit and the Robert Wood Johnson Clinical Scholars Program, Yale University School of Medicine, New Haven, Connecticut, and the Cooperative Studies Program Coordinating Center, Veterans Administration Hospital, West Haven, Connecticut. This work was supported in part by the Robert Wood Johnson Foundation Grant 6309 and a grant from the Andrew W. Mellon Foundation. Dr. Esdaile was supported in part by the Arthritis Society of Canada, the Frank McGill Fund, and Fonds de la Recherche en Santé du Québec. Requests for reprints should be addressed to Dr. Alvan R. Feinstein, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510. Manuscript submitted January 3, 1986, and accepted February 11, 1986.
*Current address: McGill University, Montreal, Quebec, Canada.

community population of 100,000 persons and find 400 cases of a particular cancer, its prevalence is 400 of 100,000 or 0.004. During the next year, if 200 new cases of that cancer develop, its incidence would usually be calculated as 200 of 100,000 or 0.002. If 300 persons of the population die during the year, the incidence of mortality, i.e., the total mortality rate, is 300 of 100,000 or 0.003. If 50 of the deaths are attributed to the cancer, the incidence of cause-specific mortality for the cancer is 50 of 100,000 = 0.0005.

Although these definitions are easy to state and illustrate, the data needed for accurate calculations of incidence and prevalence are not easy to obtain.

SCIENTIFIC PROCESSES IN ACQUISITION OF DATA

Rates of occurrence are calculated when an appropriate numerator is divided by an appropriate denominator. The statistical activities involve the calculation itself, and the various adjustments and correlations that are made with the results thereafter. The scientific activities involve the basic processes that collect appropriate information to be counted as denominators and numerators. The processes requireat least four scientific procedures in identification, surveillance, registration, and linkage. The first Wee procedures most be carried out to find and count the individual persons who will constitute the denominators and numerators that appear as the separate numerical constituents of any occurrence rate. The linkage procedure allows the denominators and numerators to be joined fw the actual calculation.

Identification. The identification of the persons who are counted in a rate requires a choice of two conditions—one for the denominator state and the other for the numerator state—and the performance of examinations to detect the presence or absence of the conditions.

For the denominators of most epidemiologic rates, the selected condition is simple. It is the existence of a person who resides in a specified geographic region. In many occurrence rates, however, the persons in the denominator are further subdivided for additional demographic attributes, such as gender, age, race, occupation, or socioeconomic status. The geographic and other demographic attributes are usually determined from information collected in the household surveys performed during the decennial census.

For the numerators, the rates depend on some other condition, which is often a particular medical event, such as development of a selected disease or death. Although the occurrence of death is obvious, the detection of a particular disease is often quite difficult, requiring special activities in clinical examination, technologic testing, and diagnostic reasoning.

Surveillance. For the denominator and numerator conditions to be appropriately identified, the persons who have those conditions must be suitably examined. The surveillance process contains the arrangements made for the examinations to be conducted with frequency, intensity, scope, and accuracy.

The persons in the denominator are usually [illegible] from the surveillance of a national census, which ducted every 10 years to collect data for every every region. The scope of the process will be inco if all the persons who live in a particular region located and counted, and accuracy will be imp individual persons offer erroneous information to t sus taker.

Because the census is conducted only once ev years, whereas rates of disease occurrence are calculated each year, the denominators must be ed in each region for the intercensal years. The strategies used for these estimations depend on data about deaths, births, and migration in the region about deaths and births are easy to get, but patte migration are often difficult to quantify in a highly society. When checked at the end of each decad intercensal estimates are sometimes found to be accurate, but in many instances, the numbers are giously wrong. Doll and Peto [2] have described so the problems that make "the available population mates. . .surprisingly unreliable," with one of the examples being the 50 percent increase, despite a ently suitable "adjustment," between the number of ed States men aged 85 years or older who were est ed in 1959, and the number later counted in the cens 1960.

Regardless of the quality with which the denomina are counted or estimated, the major scientific proble surveillance is detection of the numerator event. If event is death, its occurrence is readily identified; events such as disease cannot be confidently noted present or absent unless suitable diagnostic examinati have been imposed. Because most persons do not ceive routine medical surveillance, and because even patients who have periodic check-ups may not receive of the examinations needed to diagnose or rule out presence of every disease, the ailments identified citation as numerator events depend on diverse clini phenomena and patterns of medical practice. In the sence of a standardized medical surveillance proce substantial variations in surveillance can occur amo different persons in the same geographic region, among different geographic regions. The effects of th variations will be discussed later.

When the medical surveillance process is conduc for numerator events, the existence of a prevalent con tion can be determined at the same time that the denomi nator condition is noted. For example, in special nonce sus studies, the prevalence of certain diseases may found when persons are asked during household surve to state their particular clinical conditions or diagnoses

the numerator event is an incident *condition*, however, its occurrence cannot be accurately determined without two special acts of surveillance. In the first *act*, the subjects must be suitably examined at the *start* of the follow-up interval to find those who have already had the event, and to remove them from the *group* "at risk" in the denominator. In the second *act* of surveillance, which transcends the *mere* performance of a single diagnostic examination, the denominator persons who are free *of the* event of interest must receive the appropriate follow-up examinations that will *either* detect the new event when it appears in the future, or demonstrate *that* it has remained absent.

Registration. The information acquired during the processes of identification and surveillance does *not* become statistics until the data have been registered *at* a *suitable* site. For denominator data, this site is usually the national census bureau; but different agencies are used for the numerator data. The agencies include municipal and state health departments, as well as national bureaus of "vital statistics." In the United States, the *municipal* and *state* health departments may analyze their information *separately*, but they also transmit their data to the National *center* for Health Statistics, which then *carries out* national and regional analyses.

In the past few decades, *new* mechanisms have been developed for the registration *of* numerator data. Special agencies, such as tumor registries, have been *established* and maintained to collect routine information about certain diseases, and certain ad hoc registries have been created for special research purposes. The ad hoc registries, as exemplified by the SEER (Surveillance, Epidemiology, and End Results) system [3] for assembling data about cancer, consist of consortiums of laboratories, hospitals, or *other* medical settings in nine regions of the United States and Puerto Rico. Each consortium collects data on all new cases of cancer reported for persons in the corresponding "catchment" region.

When the registered data are assembled, the *statistical* processes begin. By counting what has been registered for the denominator, we find N, *the* size *of* the denominator group. By counting the registration of numerator events, we learn their frequency, f. These two sets *of* counts produce the occurrence rate, calculated as f/N. If *the* numerator event is a "cross-sectional" condition, determined at the same time as the denominator, f/N indicates prevalence. If the numerator *event* is "longitudinal" and noted *at* a later *date*, f/N indicates incidence.

Linkage. In the foregoing calculation, f and N were joined by a virgule—the slash mark that serves both to separate and link the two numbers. If *each* denominator person has received suitable surveillance, the occurrence (or nonoccurrence) of the numerator *condition* will be determined directly, and it *can* be easily linked for that person. If appropriate surveillance is not established, and *if* data are obtained in *some* other manner, the persons in the numerator must be suitably linked to their status in the denominator.

This linkage is not easy for most epidemiologic statistics, because the incidence rates of diseases or death are not obtained by direct longitudinal follow-up of the individual subjects counted in the denominator. The denominator information is collected at the census bureau, but the numerator information is collected at some other site. The data are joined by counting the corresponding values of N and f for persons who apparently live in the same region, but whose actual status has not been verified in both denominator and numerator. Thus, although the virgule in f/N of an occurrence rate implies that the f and N values arise from suitable processes of identification, surveillance, and linkage, these processes are often inadequate, particularly for the *incidence* calculations with which the rest of this essay is concerned.

EVOLUTION OF INCIDENCE RATES

In the early development of epidemiologic statistics during the 18th and 19th centuries, the incidence data consisted of rates of total mortality for a geographic region, such as a city, county, or nation. With this focus, almost all of the required scientific processes were relatively easy to carry out. The denominator group was simply defined as the inhabitants living in a particular region; the numerator condition was a death, regardless of cause, in any member of that group; the occurrence of death itself could readily be determined; and, with the relatively stable, nonmobile life styles that existed before the 20th century, the linkage problem was trivial. The persons who died in a particular geographic region had almost all been born there and continued to live in that region thereafter.

The main scientific challenges of the early era were in the requirements for complete registration. An effective census mechanism was needed for an accurate count of the persons contained in the denominator. For an accurate count of deaths as numerator events, each person who died had to be cited, according to law, in a death certificate submitted for registration at an appropriate agency.

The arrangement and development of these registration processes — particularly for the relatively new work of death registration — were fundamental activities of such pioneer figures in viral statistics as John Graunt, William Farr, and Lemuel Shattuck. The data obtained with the death and census registrations could provide the necessary values of f and N; and the results could easily be linked to calculate f/N as an annual mortality rate in each selected region.

An important distinction of the f/N quotient is that it is structured as a proportion, but the result was called a *rate* rather than a *proportion* (or *percentage*), because the result came from two different enumerations, using two different sources of data. The individual members of the

denominator group had not received the individual surveillance necessary to identify and enumerate each person's fate in a manner that would let f/N truly represent a proportion.

Although the annual mortality rates for a general population are really estimates rather than directly counted observations, the improved registration processes of the 19th century made the information trustworthy enough to be used for three major types of applications:

1. By comparing the death rates with simultaneously registered birth rates, demographers could plot annual trends of the population's composition in different regions. This information could be used for intercensal estimates, but was particularly important for the "statists" who studied various aspects of a nation's "political arithmetic" [4].

2. By noting the mortality rates in different age groups, actuaries could construct the "life tables" that would denote remaining life expectancy at different ages. This information could be used both demographically as an index of national "health" and also commercially to establish appropriate costs for life insurance policies.

3. By linking the occupations and socio-economic status of both the numerator and denominator groups, social-minded epidemiologists or demographers could demonstrate disparities in death rates for certain occupations and social classes.

Although the registered death certificates had a citation of the lethal disease(s), the diagnostic information in the 18th and 19th centuries was generally too unstandardized and inconsistent for effective routine usage. After efforts were made in the second half of the 19th century [5] to standardize the nomenclature of disease and to select a single disease that would be regarded as *the* cause of death, the stage became set for 20th century statistics. The "modern" statistics contain not only the total mortality rates, but also calculations of cause-specific mortality rates for individual diseases.

ADVANCES IN THE 20TH CENTURY

During the 20th century, as the registration processes became well established throughout technologically well-developed countries, several major expansions took place in the use of epidemiologic statistics for incidence:

1. The lethal causes cited on death certificates could demonstrate cause-specific mortality rates, but the data became used to show the actual incidence of the cited diseases. These rates of incidence in different demographic groups then became converted into probabilities that represented the "risk" for development of each disease in individual members of each group.

2. Changes in the annual trend of these incidence rates or risks were associated with such concomitant phenomena as environmental pollution, chemical composition of water, industrialization, and public health or medical interventions.

3. The statistical trends and associations the major beliefs about the role of these phenom causes of disease, and to major decisions in public for health and medical research. Thus, the rising trends in the annual incidence of cancer have pr the belief that a "cancer epidemic" has occu technologically advanced societies; and associat these trends with various forms of urbanization, ind ization, and other types of information have led to t that most cancers are produced by carcinogens in food, tobacco, chemicals, and other suspected s outside the human body.

SCIENTIFIC PROBLEMS IN 20TH CENTURY DAT

Although a complete scope of registration for deaths, and regional populations was generally w tablished in technologically well-developed countrie ing the 20th century, the concomitant advances in nology began to produce or demonstrate substantia scientific problems in identification, surveillance, re tion, and linkage. The identification of disease has altered by nosologic changes in concepts of diagr and by dramatic changes in diagnostic and therap technology. The frequency and intensity of surveill have been greatly increased by improved access to m cal care, and by routine examinations performe screening, employment, insurance, or other reasons. recording and analysis of registered data have been stantially enhanced by computers and other advanc data processing, and the linkage process has been c plicated by the migrations encouraged by change occupations, transportation, and sociocultural pattern

Some of the new problems affect denominators, ers affect numerators, and all of them affect the statis calculated as incidence rates for cancer.

Identification of Denominators. With incidence ra being used to show the "risk" of development of can diseases, the denominator groups could no longer easily assembled from census counts. The groups wo have to be modified to include only those persons w were actually at risk. Persons would have to be remov from the at-risk group if they did not have the particu anatomic organ in which the disease might develop, o the disease had already occurred. For example, wom cannot be counted as being at risk for the incidence endometrial cancer if they have had a hysterectomy o the cancer is already present as a prevalent condition.

The problem of absent anatomic structures was cre ed because new technology allowed various diseases be treated with a relatively safe and simple surgical re moval of a person's uterus, thyroid, gallbladder, or oth organs. The removed structures were then unavailable be at risk for the development of cancer (or other subs quent diseases). This problem, when suitably conside has usually been managed with sample surveys or o

nology, however, many cancers have been found to be *either* nonfatal or treated so successfully that they *are* not causes of death. Consequently, the incidence rates will inevitably *be* too small if the *numerators* depend solely on death-certificate citations.

Although used for more than a century, the peculiar scientific *strategy* of substituting cited causes of death for the actual occurrence of diseases in living persons has led to *many* outstanding errors. One *obvious* error, to be discussed in a later *section.* is that the rates of incidence will rise fallaciously if the count of occurrences expands to include cancers *detected* in living patients.

Choice of diagnostic titles for disease: Before a standardized nomenclature was developed in the *second* half of the 19th century [5], doctors could use a huge variety of terms as diagnostic names of disease. The development and decennial revision of a standard international system of nosography have provided a specific set of entities for use in nomenclature, but the taxonomy contains more than a thousand entities [9]. Although the diagnostic names of *these* entities have been *standardized* by international agreement, no *criteria* have ever been offered for making the diagnoses. Consequently, massive observer variability *can* occur when the diagnostic terms *are* applied by individual clinical practitioners. In *one* study [10], for example, it was noted that the same clinical condition was often called *emphysema* in the United States and *chronic bronchitis* in the United Kingdom. In a different investigation, the diagnostic terms used by practitioners seemed to depend on the medical school they had attended and their *date of graduation* [11].

In the absence of *standardized* criteria for diagnoses, two patients with the *same* ailment *can* receive *different* diagnoses from different clinicians. Nevertheless, the inconsistent diagnoses assembled in death-certificate data have constantly been tabulated as though they represented diseases.

Ascribing death to a single cause: Another consequence of modern technology and increased life expectancy has been the identification of multiple diseases in individual patients. With modern methods of diagnosis readily available and frequently employed, most patients *are* regularly found to have several diseases. many *of* which might have *acted* individually *or* collectively as a *cause* of death. Substantial difficulties and inconsistencies *are thereby created,* even after postmortem *examinations.* in deciding which disease among several was uniquely responsible for a *person's* death. Nevertheless, when incidence rates are calculated from cause-specific *mortality* data, *each* death has been ascribed to a *single* disease.

Revision of rubrics and hierarchies in coding: Because so *many* different diseases can be cited as individual causes *of* death. *statistical* tabulations *of* incidence rates for each disease would be difficult to analyze and interpret. *consequently,* groups of diseases *are* regularly consolidated into a smaller number of coding rubri are used for the ultimate enumerations. Thus, the nu *respiratory cancer* may include primary cancers trachea, bronchi, and pulmonary parenchyma. The of *digestive-tract cancer* may include primary ca located anywhere in the gut from the pharynx to the as well as cancers of the pancreas, hepatobiliary sy and retroperitoneal space. One scientific probl these coding rubrics is that their contents are reg changed in different eras, so that *pulmonary edema* be classified as a pulmonary disease in one era, and cardiac disease in another [10].

For incidence data about cancers, the greatest sci ic problem in the coding rubrics is their use of an arb set of hierarchical rankings for selecting a single cau death. Regardless of the diligence or accuracy with w a clinician may have assembled the medical evide formulated the diagnostic reasoning, and prepared death certificate, the diagnoses entered on the certi are often rearranged when the information is coded a registration agency. The agencies maintain a list of se tion rules, sometimes called "Nosology Guidelines," that provide an arbitrary set of hierarchical ranks of le ity for different diseases. Because these rules are app to all the diseases listed on the death certificate, reg less of the clinician's cited sequence of causes, agency's hierarchical guidelines may often select som thing that alters the clinician's choice of *the* cause death.

The hierarchical rules differ in different countries are regularly changed within the same country. Co quently, disease A may be regarded as more lethal t disease B in one era and nation; and their rankings may reversed at some other time and place. In the two repe ed occasions [13,14] in which the hierarchical selec process was tested, the disagreements in assigning single cause of death were so great among internati agencies that the authors of one study recommended "a new basis needs to be developed for mortality sta tics." The recommendation has not been carried out.

The arbitrary ranks of the hierarchical system are p ticularly important for the incidence of cancer, beca lethal priority is often given to a cancer whenever listed anywhere on a death certificate. For example, clinician's original decision was rejected, and cancer the lung was coded as the cause of death in a pe whose death certificate suggested that he died of myoc dial infarction seven years after successful pneumonec my [12].

Changes in nosology and technology: The scien distress that might be produced by these many caprici relatively minimal, however, compared with the impa changing nosologic concepts and diagnostic technol With new nosologic concepts of labeling disease, ce entities — such as *dropsy* or *chlorosis* — may disap simply because they are no longer regarded as disea

ing rubrics th
is, the rubric
cancers of th
ma. The rubri
mary cancer
ix to the anus
biliary system
c problem is
are regularly
y edema may
era, and as a

atest scientif
of an arbitrary
ngle cause of
cy with which
cal evidence,
prepared the
he certificate
coded at the
list of selec-
elines," [12]
nks of lethal-
s are applied
cate, regard-
causes, the
elect some-
ne cause of

ountries and
itry. Conse-
lethal than
ings may be
two report-
al selection
assigning a
ternational
nended that
ality statis-
ied out.
m are par-
r, because
never it is
ample, the
cancer of
a patient
of myocar-
mmonecto-

scientific
aprices is
impact of
chnology.
e, certain
disappear
diseases

ips]. With new methods of detecting disease, certain entities—such as systemic lupus erythematosus and diverse cancers — may appear to increase only because they now become identified during life.

If the new diagnostic methods also lead to stricter nosologic standards and criteria, certain diseases may appear to decrease because the available evidence does not satisfy the strict diagnostic criteria. For example, regardless of whatever else has contributed to the declining incidence rates of pulmonary tuberculosis in the 20th century, one prominent cause has been the demand for suitable microbiologic and roentgenographic evidence. With this demand, patients having only the clinical manifestations of fever, hemoptysis, and suitable types of pulmonary rales would no longer receive the diagnosis of tuberculosis that they would have been given almost routinely 80 years ago. Similarly, a patient dying with cachexia and an abdominal mass in whom, formerly, stomach cancer would be diagnosed, could no longer receive that diagnosis when the nosologic criteria began to demand suitable roentgenographic evidence of the cancer [10].

Despite the obvious impact of changing nosology and improved diagnostic technology, their roles are constantly ignored in the rise or fall of incidence for many other diseases, particularly cancer. The problems are further discussed in the next section.

Effects of Surveillance. Since special diagnostic technology is required to show that certain diseases are present or absent, the application of this technology is a crucial feature in the surveillance process for both prevalence in denominators and incidence in numerators. For the population under study to receive suitable surveillance, the mere invention and existence of a technologic procedure are not enough; the procedure must be disseminated and widely used.

For example, although roentgen rays and microbial smears had been developed in the 1890s, the diagnosis of pulmonary tuberculosis in 1920 at New Haven Hospital still depended mainly on a clinical history and physical findings; only 50 percent of the patients had undergone chest radiography and only 10 percent had had the tubercle bacillus microscopically demonstrated [16]. The cytologic smear used to diagnose cancer in desquamated cells was invented by Papanicolaou in 1945; yet almost a decade later, in 1954, the Pap smear of sputum was not yet being used routinely at New Haven Hospital to help identify or rule out the existence of lung cancer [10]. In the past decade, the sharply rising annual rates of pancreatic cancer have aroused suggestions that another new cancer "epidemic" is occurring, but little or no attention has been given to the role of dramatic new technology—abdominal ultrasonography, computer-assisted tomography, and endoscopic retrograde cholecystpancreatography—in detecting pancreatic cancers that formerly would have escaped identification during life.

Without adequate attention to the incidence and prevalence of surveillance with diagnostic technology, and without suitable consideration of the technologic effects on what becomes detected or left unidentified, the rates of reported diseases simply represent the rates of diagnostic detection. They cannot be accepted as scientific enumerations of the true incidence and prevalence of those diseases.

Problems in Registration. Two of the main problems in the census-bureau registration of denominators were briefly outlined earlier. For one of these problems, when the intercensal estimates of denominators are found to be wrong, the census bureau can readily correct its data retroactively. If the denominator estimates have already been used, however, to calculate myriads of annual incidence rates for an enormous number of diverse geographic regions, the attempts to create retroactive corrections would be formidable. Accordingly, the originally calculated rates are usually allowed to stand, with the fervent hope that the errors will not be too striking or misleading.

The second problem, which is more substantial, occurs when the scope of the census is inadequate, leading to the under-enumeration of certain ethnic groups. Although the consequences are most often discussed as a political issue in legislative representation, the statistical impact can be striking for rates of disease. For example, after the 1970 census in the United States; subsequent studies showed that about 10 percent of black men had not been counted [17]. If we suppose that all deaths from cancer in black men are identified and counted as numerator events in the medical settings in which the patients die, but that only 90 percent of the black men are actually counted in the census denominator, the incidence of cancer deaths, when f is divided by 0.90N rather than by N, will be 1.11 higher than it should be. The rate of cancer deaths in black men will be falsely raised by 11 percent.

Regardless of the size of the incidence-rate errors produced by these two problems in denominator registration, a different set of major problems has been produced by new modern methods of surveillance and registration for numerator events. In the new mechanisms described earlier, the data collected and set to a special regional registry can include information about cancers discovered during life as well as those identified at necropsy. The reports that are newly submitted to the registries each year are then used as numerators and converted to incidence rates when linked to the corresponding regional denominators estimated from census data.

One obvious result of this new registration mechanism would be a sharp increase in the incidence of cancer. The registered reports will include not only all the cancers cited in death certificates, but all the other identified cancers that are not lethal and that were formerly not counted as numerator events. Although substantially larger than before, the newly calculated incidence rates will

be scientifically untrustwo
to be altered by technologic
and prevalence, not having be
remain unadjusted in the re

Problems in Linkage. When an incidenc
lated as I/N, the original implication is t
subjects in the denominator were initially
found to be free of the ever
implication is that suitable follow-up exan
performed. with the numerator event four
members of the group, and absent in the o
bers.

Although incidence rates are interpreted
had these two implications, none of the
rates for the occurrence of cancer (or other
actually constructed in this manner. As not
linkage off and N is a purely statistical activ
from two different *sources:* the census bu
agency that collected the *numerator* inform
the N subjects in the denominator are exami
see whether they have *cancer;* and for eac
years of intercensal estimates, many of the
subjects are not actually counted. Of the f su
ed in the numerator, few (if *any*) were
checked *to see* whether they were indeed
contained in the denominator.

The policy of forming these rates by pure
linkage is a pragmatic *necessity* of epidemi
too many persons are involved for each p
suitably examined and followed. Neverthele
has several scientific hazards. *One* hazar
earlier, is *created* by frequent geographic mi
highly mobile *society.* Many of the
who die in a particular region may *not* have be
in the denominators; and *many* of the person
counted in the denominators *may* have receiv
w e and diagnosis in *some other* region, a
have moved away and died elsewhere.

A second hazard, called the *ecologic falla*
when some external agent—such as change
air, *sales of tobacco, or* tactics in health care—
ated with changes in the incidence of a particul
in the region [18]. Because the data for these as
are assembled as a total composite of everyth
ring in the region, *the* events that produce
change in totals for the external agent may oc
part of the region; but the main alterations in dis
occur in a *different* part of the region. The two
entities may then be fallaciously associat
events are linked in the composite totals. Thus,
rise in use of vitamins occurs in Albany while
the colon has a major rise in Brooklyn, the two
events will *be* associated m an ecologic falla
increasing colon cancer is correlated with increa
of vitamins for the entire state of New York.

The existence and magnitude of these two

ing rubrics th
is, the rubric
cancers of th
ma. The rubri
mary cancers
ax to the anus
biliary system
c problem is
are regularly
y edema may
era, and as a

atest scientif
of an arbitrary
ngle cause of
cy with which
cal evidence,
prepared the
he certificate
coded at the
list of selec-
elines," [12]
nks of lethal-
s are applied
cate, regard-
causes, the
elect some-
ne cause of

ountries and
try. Conse-
lethal than
ings may be
two report-
al selection
assigning a
nternational
nended that
ality statis-
ied out.
m are par-
r, because
never it is
ample, the
cancer of
a patient
of myocar-
mmonecto-

scientific
aprices is
impact of
chnology.
e, certain
disappear
diseases

With *new* methods of detecting disease, certain
ases—such as systemic lupus erythematosus and
se cancers — *may* appear to increase only because
now become identified during life.

the *new* diagnostic methods also lead to stricter
gic *standards* and *criteria,* certain diseases *may*
ppear to *decrease* because the available *evidence* does
not satisfy the strict diagnostic criteria. For example,
regardless of whatever else has contributed to the declin-
ng incidence rates of pulmonary tuberculosis in the 20th
entury, *one* prominent cause has been the demand for
uitable *microbiologic* and roentgenographic *evidence.*
With this demand, patients having only the clinical mani-
festations of fever, hemoptysis, and suitable types of
ulmonary rales would no longer receive the diagnosis of
uberculosis *that* they would have been given almost
outinely *80* years *ago.* Similarly, a patient dying with
achexia and an abdominal mass in whom, formerly,
omach *cancer* would be diagnosed, could no longer
eceive that diagnosis when the nosologic criteria began
o demand suitable roentgenographic evidence of the
ancer [10].

Despite the obvious impact of changing nosology and
mproved diagnostic technology, their roles are constantly
gnored in the rise *or* fall of incidence for *many* other
iseases, particularly cancer. The problems are further
iscussed in the next section.

Effects of Surveillance. Since special diagnostic tech-
nology is required to *show* that certain diseases are
resent *or absent,* the application of this technology is a
rucial feature in the surveillance process for both preva-
ence in denominators and incidence in numerators. For
he population under *study* to receive suitable surveil-
ance, the mere invention and existence of a technologic
rocedure are *not* enough; the procedure must be dis-
eminated and widely used.

For example, although roentgen rays and microbial
mears had been developed in the 1890s, the diagnosis
f pulmonary tuberculosis in 1920 at New Haven Hospital
still depended mainly *on* a clinical history and physical
findings; only 50 percent of the patients had undergone
chest radiography and only 10 percent had had the tuber-
cle bacillus microscopically demonstrated [16]. The cyto-
logic smear used to diagnose cancer in desquamated
cells was invented by Papanicolaou in 1945; yet almost a
decade later, in 1954, the Pap *smear* of sputum was *not*
yet being used routinely *at New* Haven Hospital to help
identify *or* rule *out* the existence of lung cancer [10]. In the
past decade, the sharply rising annual rates of pancreatic
cancer have aroused suggestions that another new can-
cer "epidemic" is occurring, but little *or* no attention has
been *given* to the role of dramatic new technology—
abdominal ultrasonography, computer-assisted tomogra-
phy, and endoscopic retrograde cholecystpancreato-
graphy—in detecting pancreatic cancers that formerly
would have *escaped* identification during life.

EX 1789

formation of the dead. The occurrence of death and the concomitant demographic data listed on the certificate can continue to be used successfully to calculate *total* mortality rates. These rates *can* still be effectively analyzed for the demographic, actuarial, and socioeconomic purposes *for* which the data have been used since the 19th century; and the accuracy of the rates can be improved with sample surveys *that show* the kinds of numerator and denominator adjustments needed for geographic migrations.

On the other hand, death certificates cannot be used as scientific instruments *for* denoting diseases or indicating lethal maladies. Almost 50 years ago, after carefully analyzing the "reliability of statistics of separate causes of death," Raymond Pearl [5] concluded that "the material is fundamentally of a dubious character." Except for deaths unequivocally attributable to *homicide or* trauma, Pearl suggested that the nosologic data recorded *on* a death certificate could not be *accepted* "unreservedly *at* their face value." Because the situation has not improved during the past half century, tabulations and analyses of *cause-specific mortality* rates should be regarded as a statistical idea whose time has passed.

Develop an Additional Death **Certificate as a Scientific Document.** The medical information listed *on* death certificates could become valuable scientific data if the death certificate were transformed into a scientific document. The transformation would require a new *system* wholly different from the ambiguous, inconsistent procedure in general use *today*. The legal and scientific needs could be met separately by preparing two death certificates, rather than *one*. The first would be a simple short form, intended legally to verify the occurrence of death and to *state* whether a coroner's inquest is required. The second document would serve *the scientific* purposes of indicating what diseases were believed present *at* the time of death, what *evidence* existed to sustain the diagnoses, and whether a single disease could be unequivocally regarded as *the* lethal agent. If the diagnostic and lethal-cause evidence was *equivocal*, it would so be noted and classified. Because *the* scientific evidence could probably not be suitably prepared in a routine manner throughout an entire nation, special registration regions could be established—analogous to regional tumor registries—where deliberate surveillance, monitoring, and supportive personnel could be used to maintain scientific quality in *the* procedure.

Improve the International Taxonomy of Disease. If scientific quality can be established in the basic *evidence* of death certificates, analogous improvements *can then* be instituted to repair the nosographic and clinical inconsistencies of *the* international taxonomy for nomenclature and classification *of* disease. *Some* obvious first steps would be to discontinue the custom of arbitrary hierarchical rankings for rearranging the diseases listed *on* death certificates, to develop better methods *for* dealing with multiple nosographic entities, and to establish operational criteria for making specific diagnoses.

Add Scientific Clinical Precision to Analysis of Registry Data. For scientific analysis, the information collected in special tumor registries should have an enlarged clinical scope and classification. Because of the problems in determining the previous status of a newly reported case, the annual results cannot be called *incidence* rates, and should receive some other name, such as *occurrence* rates or *reported* rates.

The rates should also be more precisely categorized, from data obtained by careful review of each patient's past history and the clinical manner in which each cancer was manifested and detected. Certain cancers will be detected because they produced symptoms or other manifestations that acted as the iatrotropic stimulus [6] in making the patient seek medical attention. In other patients, the cancer will be diagnosed while lanthanic [6], having produced no symptoms, or symptoms that had not become important iatrotropic complaints. In the lanthanic patients, the cancer would have escaped detection except for the special screening or other diagnostic test that was part of the general examination process.

In some instances, such as a routinely scheduled periodic x-ray or sputum Pap smear, the screening test is specifically aimed at finding a lanthanic cancer. In other instances, such as a lesion found on chest radiography performed for other reasons (e.g., to rule out tuberculosis, or to show cardiac size), the lanthanic discovery will be inadvertent. By dividing the newly reported cancers into complainant, specific lanthanic, and inadvertent lanthanic detections, we can get a better idea of the types of changes that are occurring in annual rates, and of the diagnostic patterns that are contributing to the changes.

If the terms *incident* and *prevalent* are to be maintained, they can be used more precisely if applied after a review of previous roentgenographic findings or other appropriate diagnostic information obtained before the cancer was overtly identified. If absent in previous suitable examinations, the cancer can be unequivocally classified as *incident*. If the patient had no pertinent previous examinations or complaints, and if the cancer was found during a test performed routinely for screening or other purposes, it can be unequivocally classified as *prevalent*. Suitable criteria can then be developed for an appropriate intermediate classification of cancers found in other circumstances. These improvements in classification will require careful attention to data that are now generally neglected or disdained: the particular iatrotropic stimulus that made the patient seek the medical examination that led to the eventual detection of the cancer; and the results of previous appropriate examinations of the same body structure in which the cancer has been found.

Use Old Medical Records to Determine the Incidence and Prevalence of Technologic Changes. At institutions that still maintain both old medical records and index

files for accessing the diagnoses cited in the records, appropriate records can be retrieved and reviewed to note the kind of information collected and used in diagnostic identifications of the same disease at *different eras*. (This was the type of research used to *show* the frequent absence of chest x-ray and microbial evidence when pulmonary tuberculosis was diagnosed in patients at a prominent academic hospital in 1920 [16]). The maintenance of voluminous old medical records, however, is often regarded as a major administrative nuisance at *most* institutions; and *the* records are increasingly destroyed as soon as they lose their value as legal documents. *Nevertheless*, the records have an irreplaceable scientific value because they are the only available information that can indicate what has truly happened in time as the result of changing nosologic concepts and diagnostic technology.

Make Better Epidemiologic Use of Necropsy Data. Another valuable but often overlooked source of scientific information in epidemiology is the existence of collected reports of necropsy *findings at different* eras. The necropsy results *are* particularly useful not for their anticipated revelations *about* a *cause of* death in a particular patient, but for their unexpected findings of other diseases [6,20]. The Unexpected. unsuspected diseases noted at necropsy *can* help indicate the size of the *undetected* lanthanic "reservoir" of disease that will become detected and reported as *"new"* cases when improved diagnostic technology is applied to living persons.

Since the subjects who undergo necropsy may not be a representative sample of all deaths. data to be analyzed should focus on the unexpected conditions found at *necropsy*, not on the apparent causes of death. For example. because necropsy may not be performed if lung *Cancer* has been unequivocally diagnosed during life, the relative frequency of *previously diagnosed* lung cancer at necropsy will not be an interpretable rate. On the other hand. in subjects who *died* of other causes and in whom necropsy was performed *for* other reasons, the necropsy discovery rate of coexisting but previously unsuspected lung cancer can be valuable. The subjects who underwent necropsy without a previous suspicion of lung *cancer* will provide a "sample" of the general population in whom death, due to other causes, allowed necropsy to be the "complete screening" examination that revealed the existence of the lung cancer.

If the necropsies in all these unsuspected subjects are used as a denominator, and if *the* lung cancers found as unanticipated discoveries *are* used as a numerator. the result will approximate the same type of occurrence rate now being calculated at tumor registries. If the necropsy discovery cases are removed from the tumor-registry numerators, the tumor registry results will show the occurrence rate of cancers mat *are detected* during life. The results in the unsuspected necropsy group will show the occurrence rate of cancers that *are undetected* during life. By noting the secular trends in these two sets of occurrence rates, we can get a clear idea of what is really happening in the existence, diagnosis, and reporting of cancers.

The argument is sometimes offered that necropsy discovery rates are not comparable to tumor registry rates because the tumor-registry discovery rates represent risk of cancer in a particular age group, whereas the necropsy discovery rates represent a subject's "lifetime." This argument seems to emerge from a peculiar old standard in statistical calculations. When a living 65-year-old man is found to have lung cancer in a screening examination, the result is counted not as a "lifetime risk," but as the occurrence of a lung cancer among similarly eligible 65-year-old men. If a dead 65-year-old man is found to have a lung cancer in a screening examination (i.e., necropsy), the result can analogously be counted as the occurrence of a lung cancer in the similarly eligible group of 65-year-old men.

Another possible objection to the proposed use of the necropsy discovery rate is that it should be viewed not as an occurrence rate, but as a proportionate mortality rate. In the latter rate, the denominator consists of a count of all the single causes cited for a series of deaths. The numerator contains a count of the frequency with which a particular disease occurs among the total number of causes of death. The proposed necropsy discovery rate, however, is quite different from a proportionate mortality rate, because the necropsy rate does not emerge from counting causes of death, and because several different diseases might be detected in the necropsy of a single subject. Thus, someone who dies of a stroke would have the stroke counted as the cause of death in either a cause-specific mortality rate or a proportionate mortality rate. The same subject's incidental unexpected necropsy findings of gallstones, coronary heart disease, cancer of the pancreas, and diverticulosis would each be counted as part of the necropsy discovery rates for those diseases.

None of these scientific improvements will be easy to achieve. Some of them—such as the development of operational criteria for making diagnoses—represent fundamental scientific challenges that have been neglected for more than a century. Others—such as better use of medical records and careful attention to the clinical phenomena that preceded diagnoses of disease—will require a resurrection of scientific attention to activities that were once regarded as important, but that have been allowed to languish in the efflorescence of technologic data. The others — such as a more creative epidemiologic use of necropsy information—will need a renewed respect for what the dead can teach the living, and a recognition that necropsy can be useful in epidemiologic research and not just in clinicopathologic conferences.

Can all these "radical" but old-fashioned scientific activities be restored or newly instituted? If we want to determine whether the rise of new technology has created cancers as well as detected them, the scientific contributions of old-fashioned forms of research may be our best hope.

REFERENCES

1. Thomas L: An epidemic of *apprehension.* Discover 1983; 4 78–80.
2. Doll R, Peto R: The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. JNCI 1981: 66: 1192–1308.
3. Greenberg ER, Colton T, Bagne C Measurement of cancer incidence in the United States sources and uses of data. JNCI 1982; 68: 743–750.
4. Hogben L: Statistical theory: the relationship of probability, credibility and error. London: George Allen 6 Unwin, 1957.
5. Pearl R: Introduction to medical biometry and statistia. Philadelphia: WB Saunders, 1940.
6. Feinstein AR: Clinical judgment. Baltimore: Williams 6 Wilkins, 1967. (Reprinted by Robert Krieger Co., Melbourne, Florida).
7. Bauer FW, Robbins SL: An autopsy study of cancer patients. JAMA 1972; 221: 1471–1474.
8. Suen KC, Lau LL, Yermakov V: Cancer and old *age.* An autopsy study of 3535 patients over 65 years old. *Cancer* 1974; 33: 1164–1168.
9. The International Classification of Diseases, 9th revision. clinical modification. ICD-9-CM. Ann *Arbor:* Commission *on* Professional and Hospital Activities, 1978.
10. Feinstein AR: Clinical epidemiology: II. The identification rates of disease. Ann Intern Med 1968 69: 1037–1061.
11. Sorlie P, Gold E: Artifactual decline in chronic ischemic heart disease deaths due to physician terminology preferences (abstr). Am J Epidemiol 1985; 122: 530.
12. National Center tor Health Statistics: Vital statistics instruction manual. pert-2. Cause-of-death coding for deaths occurring in 1968. Washington: U.S. Government Printing Office, 1968.
13. Dorn HF, Moriyama IM: Uses and significance of multiple causa tabulations for mortality statistics. Am J Public Health 1964; 54: 400–406.
14. Anonymous: The accuracy and comparability of death certificates. WHO Chron 1967; 21: 11–17.
15. Beeson PB: Some diseases that have disappeared. Am J Med 1980; 68: 806–811.
16. Feinstein AR, Milstone EB: Effects of altered diagnostic standards on the changing incidence of pulmonary diseases (abstr). J Clin Invest 1969; 48: 23a.
17. Siegel JS: Estimates cl coverage of the population by sex, race. and age in the 1970 census. Demography 1974 11:1–23.
18. Selvin HC: Durkheim's suicide and problems of empirical research. Am J Sociol 1958; 63: 607–619.
19. Oreglia A, Duncan RP: Health planning and the problem of the ecological fallacy. Am J Health Plan 1977; 2 1–6.
20. McFarlane MJ, Feinstein AR, Wells CK: The necropsy": unexpected detections, demographic s o b tions, and the changing rates of lung cancer (abstr). Clin Res 1985; 33: 269A.