The Numerical Method

in Therapeutic Medicine

MICHAEL S. SHIMKIN, M.D.

MEDICINE is an important but small component of the total of human activities. For good or bad, medicine must integrate with the social and economic as well as the scientific patterns of its community and its time. When medicine appears to be out of step in its ecology, critics are fast and numerous. The ridiculous sketches of physicians in Molière's comedies and the savage etchings of Goya, in which physicians appear with donkey heads, cannot be dismissed as mere caricature. They also portray, alas, the inadequacies and pomposities of medicine during their times.

In our society criticisms of certain aspects of medicine are not difficult to find. When such criticisms reach public consensus, they are voiced through hearings and legislative acts in the Congress of the United States. It is neither justified nor prudent to condemn these as uninformed politics, to be resisted at all points along the line. Rather, it is preferable to examine objectively the circumstances that lead to such manifestations and to initiate remedial measures in which the interests of the public and of the medical profession are restored to balance.

The case in point is the Federal Food, Drug, and Cosmetic Act of 1962, which attempts to correct, and perhaps to over-correct, some of the practices that stemmed from the cornucopia of chemotherapeutic agents that has deluged medicine during the past 25 years.

During the therapeutic doldrum of the thirties, useful drugs could be counted on one's fingers. Then came steroid hormones, sulfonamides, antibiotics, antihistamines, tranquilizers—each with dozens of modifications, some of structure and substance, others of packaging and trivial names. Obsolescence of drugs as an economic problem approached that of obsolescence in detergents. Untoward effects required special monographs, with such revealing titles as "Diseases of Medical Progress" (*1*).

It is not surprising, therefore, that regulations under the Federal Food, Drug, and Cosmetic Act of 1962 require more stringent evidence of safety for new drugs. Included are tests for genetic and teratogenic effects over several generations of animals in order to avoid a recurrence of the thalidomide tragedy. There are requirements of prompt reporting of side effects in patients and of prompt withdrawal of suspected drugs from the market. Usefulness as well as safety is introduced as a criterion of acceptability of new agents. There are also, unfortunately, provisions regarding labeling that

Dr. Shimkin, who is chief of cancer biology, Fels Research Institute, and professor of medicine, Temple University School of Medicine, delivered the 43d James M. Anders Lecture of the College of Physicians of Philadelphia on December 4, 1963. *The lecture is reprinted from the January 1964 issue of Transactions and Studies of the College of Physicians of Philadelphia.*

are interpreted as prohibiting interstate shipment of placebos required in double-blind tests. And the whole act has generated an increasingly undigestible bolus of reports, forms, and other paper that makes life unhappy for pharmaceutical firms, investigators, and government officials.

A fair estimate of the act lies somewhere between the view that this is another horrible example of government control over free enterprise, and the view that it will solve all problems. In any case, the medical profession is square in the middle of the issues and cannot avoid being involved. Mere opposition would represent a withdrawal syndrome that invites further restrictions. Complete acquiescence would imply a confidence in Washington approaching delusional euphoria.

It is my thesis that the act is promising evidence that scientific research methods may be finally catching up with therapeutic clinical medicine. And it is high time, too, since this race has been going on for more than a century, with the victories of the scientific method over tradition and the art being far from constant or sustained.

I would like to defend my thesis by historical events, which afford the clarity of retrospection.

Yesterday and Today

When the city of Philadelphia was visited by the scourge of yellow fever in 1793, Dr. Benjamin Rush was ready. His sleepless activities during the epidemic demonstrated his personal bravery and his complete confidence in his cure of yellow fever by bloodletting and purging. He was taken to task by a layman, William Cobbett, who in a pamphlet of 1800 insisted that during the height of Rush's activities deaths from yellow fever not only did not abate but steadily rose. His conclusion was devoid of charity, estimating Rush's cure as "one of those great discoveries which have contributed to the depopulation of the earth." Cobbett was convicted of slander and fined for his intemperance of pen (*2*). Benjamin Rush continued as the outstanding medical figure of his day, and he modestly admitted that, "Medicine is my wife and science my mistress." Years later this afforded Oliver Wendell Holmes with a seldom-neglected opportunity to turn a phrase. Said Holmes, "I do not think that the breach of the Seventh Commandment can be shown to have been of advantage to the legitimate owner of [Rush's] affections" (*3*).

Figure 1. Pierre Charles Alexandre Louis, 1787–1872

In 1835 Pierre Charles Alexandre Louis of Paris (fig. 1) published his studies (*4*) on the effect of bloodletting in pneumonia, erysipelas, and other inflammations, and showed convincingly that no benefit was attributable to bleeding. Louis combined careful observations in the clinic and the pathology laboratory with analysis of the course in patients treated by bleeding and in other patients not so treated. He wrote, "To assure ourselves of the superiority of one or other [treatment] . . . in any disease whatever . . . is doubtless to be done by enquiring if under these circumstances a greater number of individuals have been cured by one means than another" (*5*).

Louis named his approach the Numerical Method and opened the first chapter of clinical biometry, the application of statistics and other mathematical techniques to clinical problems. It has been a vigorous field, expanding the primitive method of Louis by the use of the

Comment

Time will tell, of course, whether the Kefauver-Harris Drug Amendments attempt to "over-correct" some of the practices of the past 25 years, as Dr. Michael B. Shimkin states. The Food and Drug Administration is making a sincere effort to apply the new law reasonably in a manner that will accomplish its objectives without over-correcting. We believe that it will be possible to do this successfully and will welcome the considered advice and the assistance of the medical profession in our efforts.

Dr. Shimkin indicates that some provisions of the amendments are interpreted as prohibiting interstate shipments of placebos required in double-blind tests. There have been such interpretations by those outside the Food and Drug Administration, but we have gone to some length to assure all who have inquired that this is not the interpretation of the agency administering the law.

Although Dr. Shimkin may believe that the new law has "generated an increasingly indigestible bolus of reports" and that "complete acquiescence would imply a confidence in Washington approaching delusional euphoria," we suspect, to use his words, that a fair estimate of the situation lies somewhere between this view and less emphatic ones.—GEORGE P. LARRICK, *Commissioner of Food and Drugs*

calculus of probability, tests for significance, correlation coefficients, and the formal design of investigations with randomized assignments, double-blind precautions, and sequential analysis. Some of these historical landmarks, as recorded by Singer and Underwood (*6*), are listed in the box. It is noteworthy, and not attributable to the fact that the compilers are British, that all the persons of this list are French or English and that no Americans are included. The introduction of mechanical and electronic computers, however, is a legitimate recent addition to the list.

Those advances lay lightly on clinical medicine. During our Civil War bloodletting and purging were still popular. The Surgeon General for the Union, Dr. William A. Hammond, was dismissed partly because he removed calomel and tartar emetic from the approved list of medical supplies (*7*). It would be interesting to compare the volume of blood spilled on the battlefields with the volume removed for assumed therapeutic intent by physicians.

But let us progress to today. In important medical journals there appears a full-page advertisement of a respected pharmaceutical company, of a progestational compound for the treatment of threatened abortion. A lovely girl, heavy with child, looks mistily into a blossoming tree. The trade name of the compound, in large red letters, is followed by the claim that 80 percent of threatened abortions are salvageable. There is a reference to an article documenting the claim. But there are two studies, one in Australia (*8*) and one in Texas (J. W. Goldzieher, personal communication) in which the progestational agent was used along with a parallel group of controls that received a placebo. In both investigations 80 percent of the threatened pregnancies were saved. This was also exactly the result in both placebo-treated groups.

And here is what Weatherall (*9*) concludes in his coldly objective review of tranquilizers, published in 1962. First, he finds that no more than one-third of clinical reports on tranquilizers are scientifically acceptable in that they contain some semblance of comparative controls. On the basis of designed clinical trials, he states, "In anxious patients the most successful if not the only successful drugs are barbiturates. They are clearly effective, and, unlike all the newer drugs, their toxicity is not gross and is well known."

Therapeutic Response

The utility of biometry in medicine is well illustrated in the consideration and evaluation of treatment in carcinoma and other neoplastic diseases. One of the virtues of using cancer as an example is that it has a clear, all-or-none end point of survival or nonsurvival. However, the principles are applicable over a wide range of situations in therapeutic medicine.

What is meant by therapeutic response with survival as the end point is portrayed in figure 2. Curve A represents the mortality of a normal population of a certain age, sex, region, and chronologic period comparable with that of the

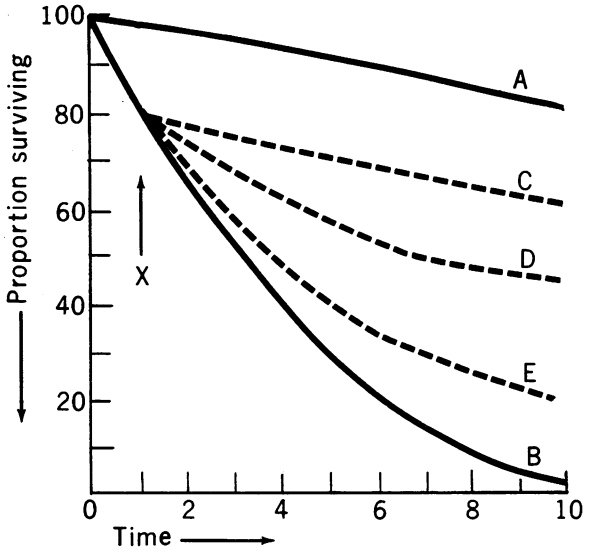
population represented in curve B, of untreated patients with a disease that accelerates the mortality rate. The difference between curve A and curve B is a measure of the additional risk of mortality incurred by the presence of the disease. Even during the time interval between 8 and 10, the slope of curve B is steeper than the slope of curve A, indicating that this increased risk is maintained throughout the total period that is represented.

If a procedure that is curative in almost all patients is carried out on the diseased population at point X, the slope of curve B would change promptly to the slope of curve C, which is parallel to curve A. This means that the additional risk incurred by the disease has been removed. Such a result might be anticipated in the treatment of in situ carcinoma of the uterine cervix. If the procedure cures an intermediate proportion of patients, there will be a more gradual deceleration of the slope, as in curve D, which remains steeper than curve A until time 6, but then becomes parallel to curve A. A therapeutic effect short of curative but yielding an objective decrease in mortality is shown by curve E, which fails to become parallel to curve A but assumes a rate significantly slower than that of curve B.

All such comparisons are valid, however, only to the extent that the groups represented by the curves to be compared are indeed comparable. Consider, for example, the effects of selection upon curve B. If treatment were instituted an interval before the point indicated by X, an apparent improvement in survival would be obtained five intervals later, despite the fact that the mortality curve is unaltered. This effect can be achieved by operating on metastatic but asymptomatic lung cancer picked up by X-ray surveys. Without actual improvement in survival, a longer treatment-to-death period would be recorded. Or, conversely, if treatment is delayed to the fourth interval, a better proportional survival would result five intervals later, simply because the mortality rate is steeper during the earlier periods. This phenomenon is involved in the paradox that in patients with cancer of the stomach or of the breast survival appears to be inversely related to the length of time between the onset of symptoms and treatment. It may also be a factor in considering patients who receive preoperative radiation if no adjustment is made for the fact that in these the operation is delayed. This delay would select out patients with the more rapidly evolving disease, with the result that a spuriously better result would be recorded among those who are operated following radiation than among those who undergo immediate surgery.

An unequal distribution of patients between the groups represented by curves E and B, by stage of disease, by the presence of other diseases or complications, by other therapeutic measures,

Figure 2. Effects of treatment on survival in a disease



Historical Landmarks in Clinical Biometry

PIERRE CHARLES ALEXANDRE LOUIS (1787–1872): The Numerical Method, 1835.

ADOLPHE QUETELET (1796–1874): Probability and normal curve, 1846.

FRANCIS GALTON (1822–1911): Correlation coefficient, 1869.

KARL PEARSON (1857–1936): Chi-square test for significance, 1900.

WILLIAM GOSSET ("Student") (1876–1937): Technique of small samples, 1908.

RONALD FISHER (1890–1962): Design of experiments, 1935.

A. BRADFORD HILL (1897–): Clinical trials, 1945.

and by a myriad of additional identifiable or unknown factors may alter entirely the conclusions that would be drawn. Comparisons may be invalid because the groups are simply noncomparable.

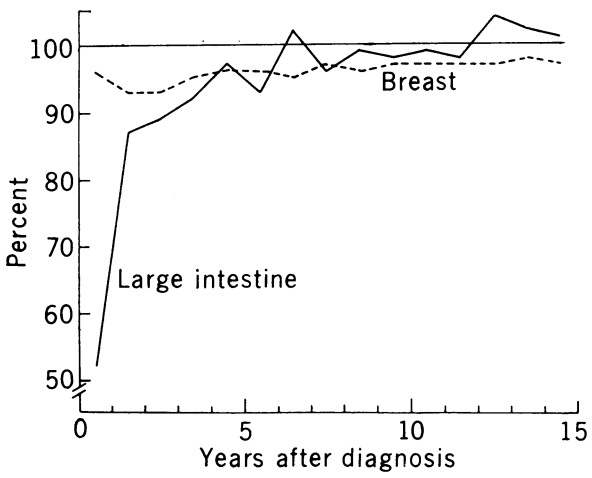
Retrospective Analysis

The statewide cancer registry of Connecticut, which has been in operation since 1935, has yielded data that bear on some of these points and also reveal the limitations of retrospective analysis.

Figure 3 shows the long-term end results in the treatment of clinically localized cancer of the breast and of the large intestine. In this graph the anticipated mortality of the population represented by curve A of figure 2 is indicated by the upper 100 percent line. The annual mortality of patients with cancer is expressed as a percentile of the anticipated mortality. When this percentile is under 100, an increased mortality risk exists; when it reaches and is maintained at 100, the increased mortality risk has been removed. This is designated as the relative mortality rate, and it is of course derived after appropriate adjustment for age, sex, and chronologic period (*10*).

In cancer of the large intestine the risk is sharply increased during the first few years, but between 6 and 7 years it reaches and remains at the level of the anticipated general mortality.

Figure 3. Relative survival of patients with localized cancer of the breast and of the large intestine, Connecticut



SOURCE: Reference 10

This is an example of curve D of figure 2. In contrast, for breast cancer, the increased risk is lower but does not reach the general mortality level during the whole observation period of 14 years. This is an example of curve E of figure 2.

The conclusions to be derived are that a proportion of patients with cancer of the large intestine treated by surgery are cured but that this effect is not demonstrable for cancer of the breast.

A recent contribution to the use of this method of analyzing survival data is the report on Hodgkin's disease and lymphosarcoma from Manchester, England (*11*), which comes to the conclusion that a proportion of these patients are cured by intensive radiation therapy. It is important additional evidence of the view that the completely pessimistic outlook that has prevailed in these diseases requires revision.

The examples given so far represent one type of retrospective analysis of data. Such analyses have a number of inherent errors, the most serious one being that the extent and nature of selection of patients for any group cannot be exactly known, so that comparisons at best must be tentatively drawn. There are techniques for retrospective pairing of patients treated by a particular method with presumably similar patients treated by other methods, and of stratifying series so that at least the obvious sources of bias are adjusted. This approach needs further exploration and has the undoubted attraction of using to full advantage experiences that are already accumulated.

Problem of Untreated Controls

It would be useful if we had reliable data on the natural survival and other features in patients with untreated disease. This is neither possible nor ethical in cancer and other serious diseases for which effective or accepted treatment is available. The "untreated" cases of cancer that are recorded in the older literature, such as Major Greenwood's classic report of 1926 (*12*), are instructive but limited as tentative baselines.

Every large group of patients with cancer today will contain a proportion that receives at least no definitive treatment. Such patients

obviously are not representative of all patients with the disease, but they should not be ignored because they do provide some information (*13*).

In many clinical problems, however, it is not strictly necessary to demand an untreated control group. After all, the desired end point is an improvement in results, so that newer procedures should be tested against the "best available" management. When there is no agreement as to what procedure is preferable, a comparison of two or more treatments is indicated. Of course, when many forms of treatment appear to yield the same results or lack thereof, suspicion should arise that none is really effective and a no-treatment group in subsequent comparisons may be acceptable. Even in such instances, palliative or symptomatic relief would be afforded to patients whenever possible, so that they would not be "untreated controls" in a laboratory animal sense. I had suggested sometime ago that the emotion-charged word, "control," be discarded in clinical investigations, and the more descriptive term, "contratest," be adopted in its stead (*14*).

Recourse is sometimes made to comparison of effects in patients who respond and others from the same series who do not respond to treatment. Almost by definition, survival, or other end point, would be superior among the responders than among the nonresponders. It may be then implied that the nonresponders represent the baseline, somewhat like an untreated group, and that the superiority of the responders is attributable to the treatment.

This situation is analagous to dividing a crate of apples into two boxes, one for large apples and the other for small apples. If the apples were of different sizes, a statistically significant difference between the two boxes would be achieved, but the implication that this sorting increases the size of the larger apples hardly would be justified. The differential between treatment responders and nonresponders likewise may be caused by a separation of favorable and unfavorable patients, without effect of treatment except as a separating device. Indeed, this may indicate that an untoward effect occurred among the nonresponders, so that the difference is due to increased mortality among them rather than a decreased mortality among the responders.

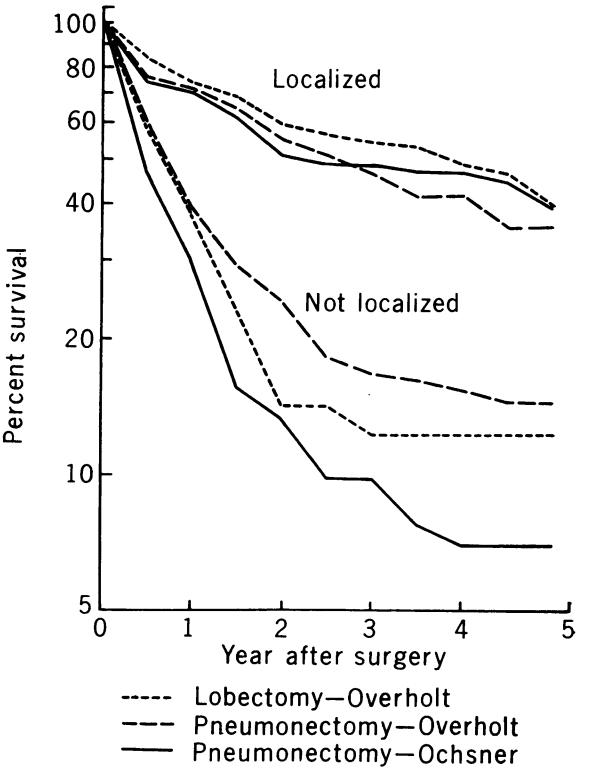
This problem is not obviated by comparisions of groups which both receive some form of treatment. An example of this occurred in an analysis of patients with lung cancer treated by pneumonectomy and by lobectomy (*15*). Among patients with more extensive but technically operable disease, survival was better following lobectomy (fig. 4). Obviously this was not because better results were achieved by the more limited operation, but because the more extensive resection increased mortality.

Designed Clinical Trials

Accepting that science is measurement and that most measurements are comparative, we come to the key question in the evaluation of clinical therapeutics. This question is: With what patients should the treated patients be compared?

Most patients are selected for treatment. Criteria for such selection differ between insti-

Figure 4. Survival of patients with lung cancer following pneumonectomy and lobectomy



SOURCE: Reference 15

tutions and between physicians. Moreover, there is known or unknown selection resulting from the type of population an institution or a physician serves. Consecutive series, historical controls, and even alternate cases have been shown not to meet the requirements of strict comparability, and such comparisons usually lead to more disputations rather than to clarifications.

The techniques of the modern designed clinical trial are applicable to these problems. One of the developers and an outstanding proponent of designed clinical trials has been A. Bradford Hill of London. I recommend for literary pleasure as well as for instruction his addresses on this subject (*16*) and his textbook on medical statistics, now in the seventh edition (*17*).

The first requirement for a designed clinical trial is the precise statement of the problem it plans to answer, the exact, specific details of the procedures it will follow, and the definitions regarding patients to be included in the study. The second requirement is that two or more groups of closely similar patients must be observed at the same time, under the same circumstances, but differing in the treatment to be compared. The third requirement is that the assignment of patients to these groups must be by some process of random allocation. And, finally, it is desirable, whenever possible, to minimize the inevitable biases of the patient and the physicians by disguising the difference in the treatments to be compared by blinding or double-blinding procedures.

The principle of randomization and the feature of blinding have been responsible for much of the resistance to, and misunderstanding about, designed clinical trials. The word "random" has been confused with "haphazard," whereas actually this is a carefully planned scheme that invokes the laws of chance in order to overcome known and unknown sources of selection and bias. The term "double-blind" has become a fertile topic of medical humor, and a staid synonym would at least reduce the mirth index.

Of course, designed clinical trials are not completed until they are analyzed, written up, and reported. In this time-consuming process the adequacy of the randomization and of all technical factors is again surveyed and may be redetermined by further "blind" interpretation of data. Statistical tests for significance and the correlation of various features of the patients to response may lead to additional dividends and the design of further investigations.

Adjuvant Cancer Chemotherapy

In 1955 the National Cancer Institute embarked upon an ambitious program toward the development of chemotherapeutic agents in cancer. Under the leadership of its clinical panel, chaired by I. S. Ravdin, a series of designed clinical trials in cancer were initiated (*18*). In addition to these drug development studies, interest was aroused in the question of whether chemicals such as nitrogen mustard, that have partial effects on some forms of advanced cancer, may have more definite effects if used at earlier stages of disease. It was accepted that such investigations were possible only if the chemicals were used as adjuvants to standard resective surgery. One of the many factors that led to the initiation of the adjuvant chemotherapy trials was that surgeons here and there were already using toxic alkylating chemicals with this idea in mind but without proof of efficiency.

Protocols were prepared for the trial of nitrogen mustard (HN_2) or triethylenethiophosphoramide (TSPA) as adjuvants for surgical resection of cancer of the lung, stomach, colorectum, and breast (*19*). At the completion of the operation, a sealed envelope from a prepared randomized set was opened. It contained instructions as to whether the particular patient was to receive the drug or was to be included in the control group. Copies of reports were sent to a central biometric office.

Two groups of investigators, one from hospitals of the Veterans Administration and the other from university departments of surgery, accepted participation in the studies.

Within a few months the Veterans Administration group reported that postoperative morbidity and mortality were increased by the use of the alkylating agents at the doses selected (*20*). In pulmonary resection, the 30-day mortality increased from 12 percent without HN_2 to 22 percent with HN_2, and in gastric resection, from 11 percent to 23 percent with TSPA. A

wide variety of complications and untoward happenings were reported with greater frequency in patients receiving the alkylating agents.

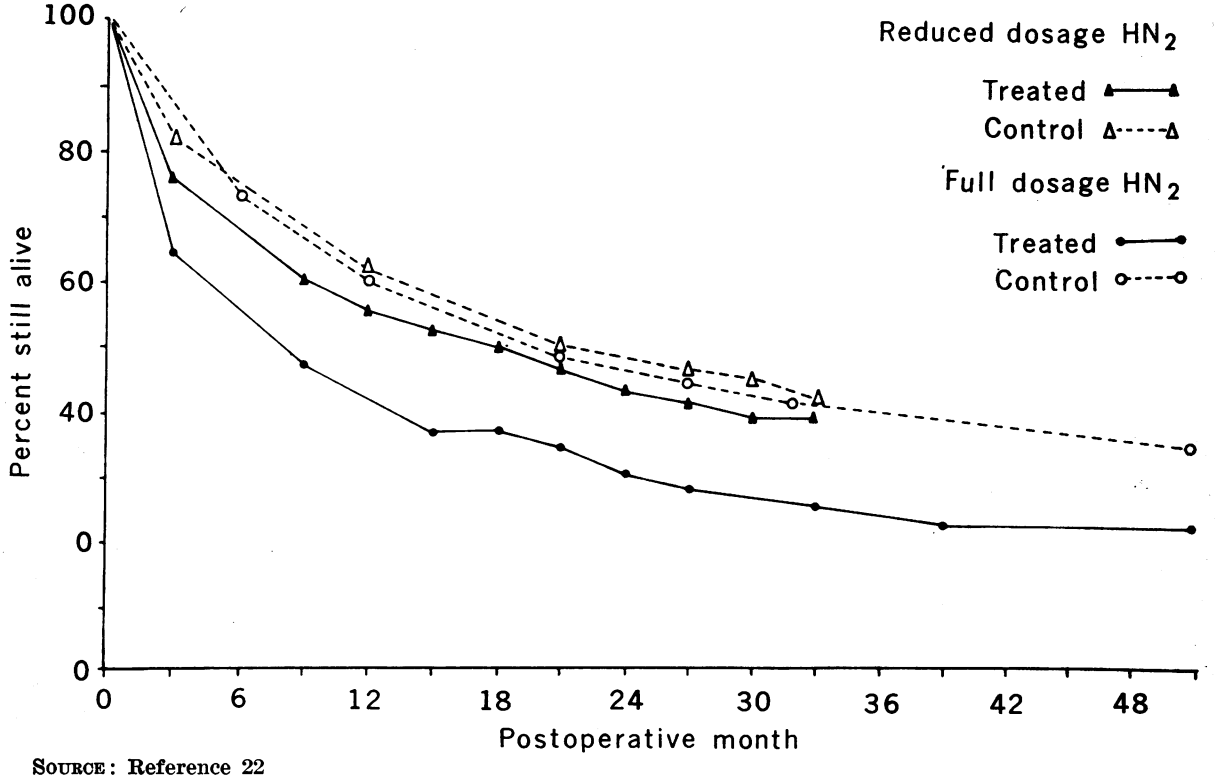
The program was continued with the substitution of two coded types of ampouled material, given postoperatively on a randomized basis and double-blinded. Both of these materials were sodium chloride. Postoperative complications promptly equilibrated between the two groups of patients, and the postoperative mortality dropped to that previously observed among the control groups.

Although the university experience failed to indicate an increased mortality with the addition of HN_2 or TSPA to surgery, the investigators decided to reduce the course of HN_2 from 0.4 to 0.3 mg. per kilogram, and TSPA from 0.8 to 0.6 mg. per kilogram, and to double-blind the study (*21*). All patients now received the contents of similar ampoules, some of which had the drug and others contained a placebo, sodium chloride. With these elaborations, no increase in postoperative mortality or complications was experienced.

The results of the Veterans Administration investigation (*22*) in cancer of the lung are summarized in figure 5. It is clear that survival was not improved by the addition of nitrogen mustard. Analysis of various subgroups failed to reveal more subtle differences, and the results of the university investigators were superimposably the same. The answer is clear; there is no indication for using HN_2 at tolerated doses following resection for lung cancer. At higher doses there is a possibility of increasing mortality. Parenthetically, the results in the adjuvant use of TSPA in cancer of the stomach and of the colo-rectum were as negative as in the lung cancer study.

A similar study of the use of TSPA in radical mastectomy by university investigators (*23*) is summarized in figure 6. Here the result is interestingly positive and shows that this adjuvant procedure does yield a demonstrable decrease in recurrences. The first phase of the investigation has been criticized because of the unusually high rate of recurrence in the control group. It is probable that during this phase many participants did not include in the study

Figure 5. Survival of patients with lung cancer with and without postoperative nitrogen mustard as an adjuvant therapy (Veterans Administration adjuvant cancer chemotherapy group)



SOURCE: Reference 22

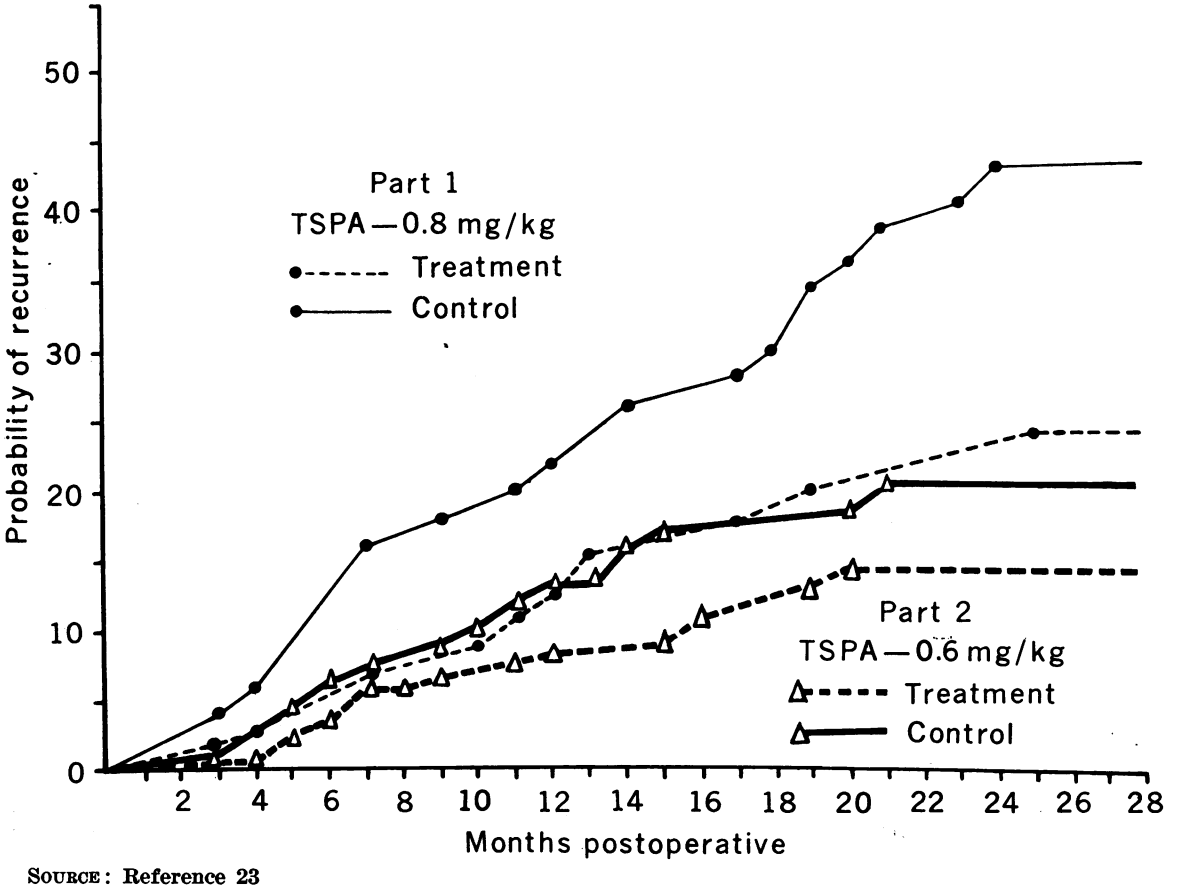
patients with smaller, localized breast cancers, so that they entered neither the treated nor the control groups. With further experience, and suggestion that an effect was being observed, the second phase of the study probably contains a more representative sample of operable mammary cancer.

Analysis of the subgroups in this investigation showed that the major effect was evident in younger, premenopausal women with metastases to the axillary lymph nodes. Their overrepresentation in the first phase of the study probably contributed to the greater differentiation between the treated and the control groups. Here again we have some definite answers, indicating that TSPA is a useful adjunct to surgery for breast cancer with regional metastases in premenopausal women. The investigations are now continuing with other experimental drugs and procedures.

The conduct of a designed clinical trial, like all complex investigations, is difficult and full of pitfalls. This is especially true when the condition that is studied has important psychological components and when the responses involve subjective reactions of patients or subjective interpretations by investigators. Walter Modell (*24*) has written informatively and wittily about some of these problems.

Designed clinical trials also have ethical considerations that require understanding. For example, is it ethical to withhold treatment from patients or not to know whether a patient is receiving a possibly active drug or an inert placebo? A. Bradford Hill (*25*) has contributed a thoughtful discussion of these and related questions, and his conclusion, in effect, is that the answers depend upon the specifics of the specific situation. I, too, have been involved in many discussions of the knotty subject of experimentation on human beings (*26*) and have come to the conclusion that in most situa-

Figure 6. Recurrence of cancer in patients with breast cancer with and without postoperative triethylenethiophosphoramide (university adjuvant cancer chemotherapy group)



SOURCE: Reference 23

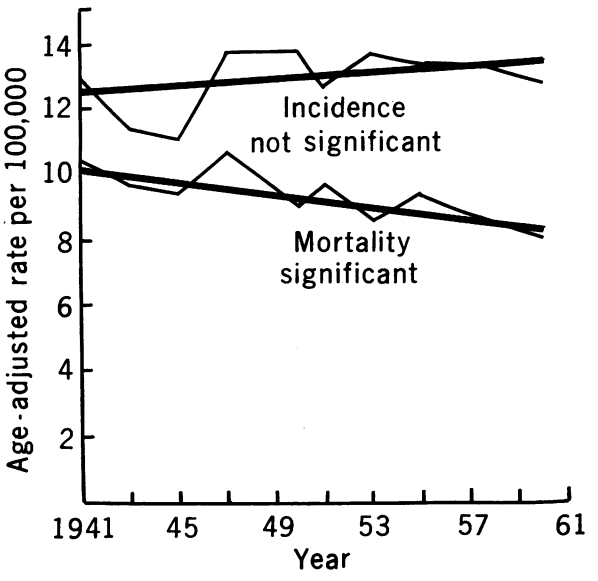
tions there are no irreconcilable differences between sound science and the ethical practices of medicine. This does not mean that there are no problems, for these will continue to exist in all considerations involving the frailties of human judgment. Particularly vexing are studies on children and other people who are legally incapable of giving their consent to participating in such investigations. The tentative answer here is that human judgment, carefully reviewed by technically qualified groups of experts, must be invoked. It should be fully realized that errors and criticisms cannot be entirely avoided, but that the price of inaction is greater than the risk that may be incurred.

Therapeutic Medicine and Public Health

The primary concern of clinical medicine is with the individual patient; the primary concern of public health is with the community of people. The goals of both are identical, for individuals are but units of a community, and a community is but a group of individuals.

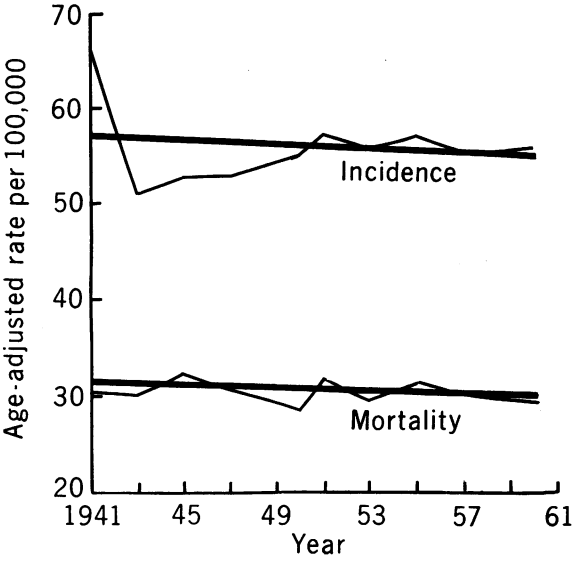
The end point of therapeutic medicine is also the effect its contribution has on all patients. A useful therapeutic modality should reflect itself in the reduction of the mortality or morbidity from the disease in the total population. Conversely, when these effects are not demonstrable, further examination of the therapeutic approach certainly seems indicated. Examples of this are afforded by the analysis of data on cancer recently published by the New York Department of Health (*27*).

Figure 7. Incidence and mortality of males with cancer of the rectum, New York State exclusive of New York City, 1941–60 (age-adjusted rates per 100,000)



SOURCE: Reference 27

Figure 8. Incidence and mortality of females with cancer of the breast, New York State, exclusive of New York City, 1941–60 (age-adjusted rates per 100,000)



SOURCE: Reference 27

Figure 7 shows the trends in cancer of the rectum in males. During the 20-year period 1941–60, there has been no significant change in the incidence of the disease; that is, the annual age-adjusted rates of occurrence of rectal cancer, per 100,000 males, have remained the same. During this period, the annual mortality from the disease has shown a modest but significant decline of about 15 percent. We cannot distinguish, from these figures alone, the effect of possibly better treatment from the influence of earlier recognition. Other studies suggest that the decline is attributable primarily to an increasingly greater proportion of patients that are surgically resected and thus, presumably, resectable. This analysis indicates that a partially effective form of therapy is having a demonstrable effect upon a defined population. Further emphasis on earlier detection and treatment is supportable by these data.

Figure 8 summarizes the trends in cancer of the breast in females. In contrast with the findings in rectal cancer, both the incidence and the mortality have remained entirely stable over the last 20 years. The difference between incidence and mortality is an index of survival, either natural or therapeutically induced. These data show that neither earlier detection, such as was hoped from self-examination of the breast, nor changes in the treatment of breast cancer during the past 20 years have been reflected in any demonstrable effect in the population. The results would suggest a careful reappraisal of our concepts regarding breast cancer and of the extent to which these concepts are being realized (*28*).

We have now taken the full tour, from Louis in 1835 enunciating his Numerical Method, to the elaborations and developments that have occurred in clinical biometry during the subsequent century. It is a difficult, time-consuming process to establish the efficacy and the indications of any form of treatment, calling for careful planning, meticulous detail, group participation, and conservative interpretation. It also means more work, because the completion of any study is but another step in a never-ending continuum. In this continuum, retrospective analyses and population comparisons also continue to be valuable sources of information.

It is evident that, in order to achieve the full development and utility of these sources of knowledge, there must be cooperation and collaboration between physicians and their institutions, with central biometric staffs to assist in planning, execution, and analysis of clinical investigations. Despite the complexities and the shortcomings, these methods are the most economical and the most reliable ones at our disposal, through which we can derive answers to many problems of therapeutic and preventive medicine. And these answers may treat roughly some cherished traditions and preconceived notions which we have long held so dear.

The words of Louis (5) are as relevant today as in 1835: "The only reproach which can be made to the Numerical Method . . . is that it offers real difficulties in its execution. . . . But . . . research of truth requires much labour, and is beset with difficulty." As physicians, we can fulfill our role by contributing to and accepting scientific quantitation rather than hallowed tradition. It is the only way we can also fulfill the valued admonition of Hippocrates, "Primum non nocere." Freely translated into the English language of today, this is "Before you are useful, be sure you are harmless."

REFERENCES

(*1*) Moser, R. H.: Diseases of medical progress. Charles C. Thomas, Springfield, Ill., 1959.

(*2*) Shryock, R. H.: The development of modern medicine. A. A. Knopf, New York, 1947, pp. 138–140.

(*3*) Garrison, F. H.: An introduction to the history of medicine. Ed. 4. W. B. Saunders Co., Philadelphia, 1929, pp. 378–380.

(*4*) Louis, P. C. A.: Recherches sur les effets de la saignée. J. B. Bailliere, Paris, 1835.

(*5*) Louis, P. C. A.: Essay on clinical instruction. (Translated by P. Martin), S. Highley, London, 1834.

(*6*) Singer, C., and Underwood, E. A.: A short history of medicine. Oxford University Press, Inc., New York, 1962, pp. 721–727.

(*7*) Eisenschiml, G.: Medicine in the war. Civil War Times Illustrated 1: 4–7, May 1962.

(*8*) Shearman, R. P., and Garrett, W. J.: Double blind study of effect of 17-hydroxyprogesterone caproate on abortion rate. Brit Med J No. 5326: 292–295, Feb. 2, 1963.

(*9*) Weatherall, M.: Tranquilizers. Brit Med J No. 5287: 1219–1224, May 5, 1962.

(*10*) Cutler, S. J.: A dynamic approach to the measurement of therapeutic end-results. *In* Proceedings of the Fourth National Cancer Conference, J. B. Lippincott Co., Philadephia, 1960, pp. 187–193.

(*11*) Easson, E. C., and Russell, M. H.: The cure of Hodgkin's disease. Brit Med J No. 5347: 1704–1707, June 29, 1963.

(*12*) Greenwood, M.: A report on the natural duration of cancer. Ministry of Health Reports on Public Health and Medical Subjects, No. 33, London, 1926.

(*13*) Shimkin, M. D., Griswold, M. N., and Cutler, S. J.: Survival in untreated and treated cancer. Ann Intern Med 45: 255–267 (1956).

(*14*) Shimkin, M. D.: Natural history of neoplastic disease. *In* Physiopathology of cancer, edited by F. Homburger. Ed. 2. Hoeber-Harper, New York, 1959, pp. 855–871.

(*15*) Shimkin, M. B.: Connelly, R. R., Marcus, S. C., and Cutler, S. J.: Pneumonectomy and lobectomy in bronchogenic carcinoma; a comparison of end-results of the Overholt and Ochsner clinics. J Thorac Cardiov Surg 44: 503–519 (1962).

(16) Hill, A. B.: Statistical methods in clinical and preventive medicine. Oxford University Press, Inc., New York, 1962.
(17) Hill, A. B.: Principles of medical statistics. Ed. 7. Oxford University Press, Inc., New York, 1961.
(18) Zubrod, C. G., et al.: Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. J Chronic Dis 11: 7–33 (1960).
(19) Shimkin, M. B., and Moore, G. E.: Adjuvant use of chemotherapy in the surgical treatment of cancer. JAMA 167: 1710–1714 (1958).
(20) Veterans Administration Adjuvant Cancer Chemotherapy Cooperative Group: Status of adjuvant cancer chemotherapy. Arch Surg 82: 466–473 (1961).
(21) Moore, G. E., Ross, C. A., and Stiver, R. B.: Chemotherapy as an adjuvant to surgery. Amer J Surg 105: 591–597 (1963).
(22) Hughes, F. A., and Higgins, C.: Veterans Administration surgical adjuvant lung cancer chemotherapy study: present status. J Thorac Cardiov Surg 44: 295–308, September 1962.
(23) Noer, R. J., et al.: Breast adjuvant chemotherapy: effectiveness of Thio-Tepa (Triethylene Thiophosphoramide) as adjuvant to radical mastectomy for breast cancer. Ann Surg 154: 629–647 (1961).
(24) Modell, W.: The protean control of clinical pharmacology. Clin Pharmacol Ther 4: 371–380 (1963).
(25) Hill, A. B.: Medical ethics and controlled trials. Brit Med J No. 5337: 1043–1049, Apr. 20, 1963.
(26) Shimkin, M. B., Guttentag, D. E., Kidd, A. M., and Johnson, W. H.: The problem of experimentation on human beings. Science 117: 205–215 (1953).
(27) Ferber, B., Handy, V. H., Gerhardt, P. R., and Solomon, M.: Cancer in New York State, exclusive of New York City, 1941–1960. Bureau of Cancer Control, New York State Department of Health, Albany, 1962.
(28) Shimkin, M. B.: Cancer of the breast. Some old facts and new prospectives. JAMA 183: 358–361, Feb. 2, 1963.

To Study Structural Components of Bones and Teeth

A new research program at the Hospital for Special Surgery, New York City, will study basic structural components of the bones and teeth as part of an investigation on diseases of these hard tissues in aging. In announcing a 7-year grant from the National Institute of Dental Research for the study, Dr. Luther L. Terry, Surgeon General of the Public Health Service, pointed out that further knowledge of basic hard tissue structure "has direct application to clinical problems." Scientists hope that future treatment of "hard tissue diseases" may be developed from such knowledge.

These diseases strike a large number of our population in middle and old age. Arthritis and rheumatism afflict nearly 10 million persons over 45 years old. Periodontal disease is the greatest single cause of tooth loss after age 35. Osteoporosis, effecting collapse of vertebrae and weakening of the bones, is a common bone disease in the middle aged and elderly.

Under Dr. Aaron S. Posner, associate director of research at the hospital and an associate professor at Cornell Medical School, a team of physicists, chemists, and crystallographers, as well as dentists and physicians, will use new and sophisticated equipment including X-ray diffraction, electron diffraction, electron microscopy, and low temperature nitrogen adsorption in the course of their program. They will try to determine why some of the crystals in bone and tooth enamel are arranged in a special way while in other bones and in the tooth dentin the crystals are randomly oriented. They will seek to understand what the link is, if any, between the mineral and protein in hard tissue. The investigators will study the arrangements of atoms and molecules which make up the proteins and minerals in bones and teeth and identify the chemical changes which take place under normal and disease conditions. Under both conditions they will investigate the mechanism of bone and tooth formation.