Who Took the Clinical Out of Clinical Research?—Mouse versus Man: Seventh David A. Karnofsky Memorial Lecture—1976

Emil J Freireich
University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

When I learned in 1976 that I had been chosen as the seventh member of the American Society for Clinical Oncology to deliver the prestigious Karnofsky Lecture, I was not only honored, but inspired. I felt that it would be my responsibility to discuss a topic that was of great importance and might have an impact on the future of clinical cancer research.

The organizing committee for this Supplement recommended that the transcript of that talk be reproduced in conjunction with this publication to contrast my 1976 views with the views expressed at the time of the Festschrift Symposium in 1997. Moreover, the committee has offered me the opportunity to augment this publication with these editorial comments.

In 1976, the reaction to my Karnofsky lecture was conservatively characterized as "highly controversial." There was a nucleus of innovative scientists who reacted positively to the talk, finding that it was consistent with many of their own experiences. A minority of fellows who reacted positively to the lecture carried some of "Freireich's Laws" into their careers. However, the dominant reaction was extremely negative. In fact, it could be characterized as "hostile." The American Society for Clinical Oncology did not yet have an official journal; the Journal of Clinical Oncology began its first publication in 1983. Because of this and the controversial nature of the lecture, it was never formally published.

I would like to comment on the seven topics that I discussed in the 1976 Karnofsky lecture, which are subsequently reviewed in this Festschrift issue. The first topic was Regulation. There is no doubt that the problems I identified 20 years ago have led to a continuing escalation in the time from conception to marketing of a new chemical entity and to a progressive escalation of the cost. Both of these continue to be significant impediments to clinical research and to the development of innovative treatment. The positive change in attitude that has occurred over the last 20 years is the appreciation of benefit: risk ratios. I think that largely as the result of the politically active AIDS community, the regulations have been modified so that risk prevention is balanced against the seriousness of the disease and the potential for benefit. Thus, the direction over the last 20 years has been unfavorable, but the predictions for the next millennium indicate a significant continuing improvement in the potential for reducing the impact of regulation of clinical cancer research. Clearly, for the patient with cancer who has not only a life-threatening disease but a limited prognosis, the requirements for preclinical testing can be substantially reduced. The need for continuing regulatory reform is emphasized both in the 1976 lecture and in my continuing efforts made in the academic medical community.

The Health Care Delivery crisis has now evolved into a major threat to clinical cancer research. The emphasis on delivery of the best available health care in contrast to the investment necessary to develop the knowledge to improve both the quality and the cost of that care has now led to the "managed care crisis." Clearly, there is a need for greater support for the clinical research that is so crucial to improving cancer control.

As for Priorities for Investigation, the last 20 years have seen a tremendous growth and emphasis on "quality of life." We now have a significant euthanasia community who advocates a hospice type of terminal care for patients with hopelessly untreatable cancers. In contrast, a clinical research environment adds an important ingredient to the patient's quality of life, that is "hope" to replace hopelessness. Certainly, the last 20 years have seen a major improvement in both the curability and the quality of life for many patients. This has been a result of clinical research activities, and many patients with hopeless diagnoses have benefited. Phase I and Phase II studies are crucial to developing new knowledge, and, for each patient participating, they offer the hope that the treatment will benefit them.

The section on Statistical Tyranny was, of course, amateurish, but important at the time. Over the last 20 years, statistical theory has become infinitely more sophisticated. The introduction of stopping rules, techniques for interim analysis, Bayesian statistics, and so forth has greatly improved the quantitative techniques for the conduct of clinical research. Many statisticians felt that this lecture was a critique of statistical input into clinical research. Quite the contrary. The intent was to point out that clinicians who have only a superficial understanding of probability theory and statistical concepts should be cautious in the use of statistics for guiding their judgement of the effectiveness of treatment. The best strategy, and one I have adopted throughout my career, is to rely on the expertise of an outstanding biostatistician. I have been fortunate to be able to consult with Dr. Ed Gehan and many other statisticians.

The Academic Requirement question is one that has grown in importance in the last 20 years. The physician-scientist has been progressively excluded from the RO1 pool, and, as I have emphasized in my Festschrift paper (1), it is important for us to recognize the achievements of clinical scientists by different criteria than those that have been used for laboratory-based research.

The False Positive—False Negative statement emphasizes the necessity for the clinical investigator to have an optimistic view of the outcome of clinical trials and to focus his attention on discovering newer approaches rather than attempting to disprove claims made by other investigators.

The final subject of Ethics has a timelessness. The clinical investigator, as a physician, has always had the responsibility of putting his patients' interests ahead of the interests of any other socioeconomic or political forces.

As I reread my 1976 lecture, I find myself proud of having had the courage to expose my opinions, prejudices, and ignorance to the academic medical community. After the subsequent 20 years of full-time clinical research in cancer, I find that I personally feel that Freireich's Seven Laws are a useful guide for my continuing research, and I believe that many other clinical scientists have found them to be useful formulations of concepts that they hold dear. Hopefully, publication of the 1976 Karnofsky Award Lecture will have a positive influence on clinical investigators as we move into the next millennium.
Introduction

The American Society of Clinical Oncology is a descendant and sister organization of the American Association for Cancer Research. Many of us are active participating members of that organization. We have an abiding interest in the fundamental scientific problems of cancer. We have a commitment to the scientific method and to research as an approach to this problem and to its ultimate solution. The creation of a separate society for clinical oncology 12 years ago reveals a distinction between clinical investigation and other types of research activity. I have chosen this topic for my presentation because it is my view that clinical research is currently at a crucial phase in its history. The clinical investigator is being pressed from many sides into extreme positions, with the result that our attention is frequently diverted from our real objectives. These forces serve as obstacles in the way of the clinical investigators’ efforts to effectively serve their patients, the community, and various government, academic, and administrative institutions. These obstacles have become so imposing that they threaten to choke off the significant clinical research that is essential to our ultimate goal of the control of cancer. It is my hope to identify these problems, to recognize that simple solutions are not likely to be found, and to appeal for moderation and reason to replace the strident polarized views that, in my view, interfere with progress.

My subtitle is “Mouse versus Man,” in the spirit of the former being our arena of experimentation so that we might find treatment for the latter. I intend to examine principles that have been developed in basic and laboratory research and show how such principles must be modified in the clinic. I need to emphasize that I am not talking about “clinical” in any managerial sense. I am using the word “clinical” to represent the relationship between two individuals, a patient and a physician, functioning as the best-qualified person to advise the patient about his illness. All of the technology and methods of science, if they are to serve the needs of man in controlling disease, must ultimately reach to the clinical, that is, doctor-patient situation. Whereas the clinical investigator is aware of the differences between mouse and man, occasionally, terms and procedures from the laboratory have come to have similar connotations for man. We have all heard speakers at this meeting “randomize” patients, and we have all heard speakers talk of “lethal drug toxicity.” These phrases, of course, are only words, not actions, but their common use by clinical investigators does reflect a dimming of the sharp distinctions between the knowledge obtained and the benefit obtained. I will attempt to cover in this paper the highly specific requirements for clinical research to sharpen the distinctions between mouse and man.

For each of the subjects listed in Table 1, I plan to identify the problem, to deal with its effect on continuing progress in clinical research, and to contrast the experimentalist’s approach to the clinical investigator’s approach. Finally, for each subject, a guiding principle will be formulated in the form of “Freireich’s Seven Laws.” The choice of an eponym does not lay claim either to originality or to any ultimate truth. It does serve, rather, to locate the principle in time and to assign responsibility for its statement to the author. Although Freireich’s Laws, as those individuals who know me recognize, are regularly violated, nonetheless, their formalized expression may serve as a reminder to return to a more moderate and more targeted direction. After a quick perusal of the seven topics to be covered, I hope you will find items of sufficient controversy to keep your interest and attention. To those who do find a violent disagreement, I would urge you not to despair, for on one of the other six issues, you may find an equally major agreement. I hope to open people’s minds to thoughtful, considered, intellectual deliberation so that we can move in the direction of reducing slightly these obstacles to research.

Regulation

The clinical investigator faces oppressive regulation from many sources. The most significant is the federal government’s regulations. The most serious problem is currently the United States FDA’s regulation of drug development.

The nine letters contained in Fig. 1 have occupied more of our professional time than probably any other single administrative problem. As citizens as well as patients and clinical investigators, we can fully appreciate the importance of regulation in the area of new drug development. No one wishes to return to an era in which remedies of little or no value or even of potential harm are widely available in the community. Certainly the public needs a regulatory body to protect them from dangerous, costly, and unnecessary exposure to drugs. The Food and Drug Act that created the regulation of medicinal drugs dates back to 1906. In 1938, after an incident involving a drug that caused more than 80 individuals to die, Congress amended the 1906 Act to require the submission of a NDA for approval by that agency. NDAs had to provide proof of safety in humans before medications could be made generally available. The next 20 years saw the Golden Age of drug development in our country: the development of antibiotics and a host of potent drugs for the treatment of human disease, including the discovery and marketing of the first anticancer drugs in the early 1950s. Thus, this regulation was effective in providing safe

---

Table 1

<table>
<thead>
<tr>
<th>1. Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Health care delivery</td>
</tr>
<tr>
<td>3. Priorities for investigation</td>
</tr>
<tr>
<td>4. Statistical tyranny</td>
</tr>
<tr>
<td>5. Academic requirements</td>
</tr>
<tr>
<td>6. False positive–false negative choices</td>
</tr>
<tr>
<td>7. Ethics</td>
</tr>
</tbody>
</table>

1 Presented at the 12th Annual Meeting of the American Society of Clinical Oncology, May 1976, Toronto, Ontario, Canada.
2 To whom requests for reprints should be addressed, at University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030.
3 The abbreviations used are: FDA, Food and Drug Administration; NDA, New Drug Application; IND, Investigational New Drug; UDGP, University Diabetes Group.
medication and permitting effective drug development. In 1962, after the thalidomide tragedy, the FDA legislation was amended by Congress, requiring an IND approval prior to any testing in man. Furthermore, the FDA could now specify the tests required to be completed after the IND for the development of the NDA for making the drug widely available. Moreover, in addition to proof of safety, proof of effectiveness must be established in moving from the IND to the NDA. This law resulted in a dramatic increase in the amount of preclinical research necessary to obtain an IND approval for the testing of a drug in man. It also progressively and enormously increased the amount of data necessary for the clinical investigator to produce before the drug could be generally available, that is, before a NDA is issued. The combined effect of these two activities resulted in horrendous increases in the cost of developing a new drug and a great increase in the time required to take a new drug from its first observation to its general availability. Were that not sufficiently bad, a crucial change has occurred within the last 2 years. Now it is clear that those responsible for the review of the IND have taken the position that the same stringent requirements be used for the development of any new cancer drugs as are used for headache medications or other drugs intended for relatively healthy people. It should be emphasized that this change occurred not at the level of congressional action, but within the regulatory agency itself. This decision brought the cancer drug development process to an abrupt and absolute halt for many months. The individuals responsible for interpreting the law did a perfect job of protecting normal individuals against the potential toxic effects of drugs designed for patients with cancer. Thus, the needs of the overwhelming majority of the population were fulfilled. Individuals responsible for making such decisions were promptly promoted and may even get a medal. But what of the minority of the population that we as clinical scientists must face in our offices and clinics, that is, the patient with cancer? For him to be told that the tyranny of the majority has interfered with the development of drugs that offer his only hope for cancer control is a cruel irony. We, as clinical investigators, must represent the interests of the community of individuals who have cancer. For these patients, our responsibility is the discovery and development of new and important medications that offer these individuals hope for an effective and creative life. It is unreasonable to expect that the time, expense, and effort involved in developing drugs for relatively healthy individuals should be the same as the standards applied to patients with cancer. We must be prepared to point out clearly the changes in the benefit-risk ratio associated with the progressively more serious prognosis of advancing clinical disease. At the same time that we turn to the public for financial support to assist our clinical research, we are faced with the reality that it is now estimated that the time necessary to bring a new drug to market is over 10 years, and the cost is as high as $22 million. Before the regulatory fervor results in a significant decrease or even a complete halt in drug development, we must emphasize to the public, to our government, and to the regulatory agencies the need for reason and judgement. We must attempt to shift from a regulatory process designed to keep drugs off the market to a regulatory process that recognizes the need for drug development in an area such as cancer treatment. We must not sacrifice the momentum we have gained toward developing more effective drugs for cancer control to fulfill regulatory requirements designed to protect relatively healthy people.

A specific suggestion has been made by Dr. Emil Frei III and by many other investigators. Specifically, this recommendation is to use a peer review system analogous to that developed by the NIH for grant review. Because it would be impossible to develop the necessary expertise capable of reviewing the applications for all types of drug development within a single federal agency, it should be possible to use outside experts in disciplines such as cancer treatment to review INDs and NDAs in accord with the published regulations and with the guidance of the executive secretary of the review body, who would be a FDA employee. In such a circumstance, the best, most capable, and most experienced minds in the community can be brought rapidly to bear on decisions relating to whether INDs and NDAs should be issued. Many of you will say: "The FDA does have advisory committees, and these serve excellent and important functions." However, the actual review of drug application is currently in the hands of the FDA staff, and such reviews are inordinately bureaucratic and routine.

Another area of regulation has become sufficiently oppressive so that the conduct of clinical research is threatened. This is informed consent. With our progressively increasing level of education and intelligence and with an increasing availability of medical care to all segments of our population, it becomes increasingly necessary that patients be fully informed, not only of the nature of their care, but of the nature, objectives, risks, hazards, and benefits of any clinical investigation undertaken in man. There seems little reason to doubt the wisdom or importance of this position. Each of us would want to have full and explicit information up to the limit of our ability to comprehend the advice or the recommendation being made by our physician. The problem derives from the implementation of such a program. How to assure that patients are informed and that consent is given? In a fit of bureaucratic activity, we have regulations that require virtually completely explicit, written documents for patients to review and sign in the presence of a witness before any type of clinical investigation can be undertaken. The specter of malpractice litigation looms behind the written informed consent and is ever present in the environment of the clinical investigator. Every institution is now required to have a review committee to conduct surveillance of clinical research activities on a regular basis. Clinical research proposals must be explicit, written, and of sufficient detail so that an expert committee of this type can evaluate their potential safety. Part of our preparation includes the preparation not only of a written protocol, but also of an explicit informed consent document for each experimental procedure to be undertaken. After
review by our investigational review committee, there is frequently another level of review undertaken at the funding agency. Aggressive implementation of such regulations has resulted in an extraordinary mountain of bureaucratic excess, which has itself become a significant impediment to research.

Perhaps more important than the investment of time, money, and effort into creating these elaborate, explicit informed consent documents is the profound impact that such a document has on the patient-doctor relationship. It is hard to imagine any area of human activity in which the written documentation of the potential hazards in complete and explicit detail would not cause tension and anxiety in the person about to undertake such a procedure. I have seen a cartoon characterization of a person boarding a jet airliner in which the captain is informing the potential passenger of all the possibilities for a catastrophic outcome of such a flight. I am confident that many individuals would be dissuaded from travel by such a procedure. It is clear that every area of our existence is associated with some risk. A patient, particularly a patient with a life-threatening illness, is seeking counsel from his physician, the best-qualified individual to give him such counsel. We must adopt procedures that allow the physician to transmit as much information as he possibly can to the patient.

To summarize the first of my seven topics: I would like to formulate Freireich’s Law Number Seven (Fig. 2), which in other circles is known as the Regulator’s Creed.

It is a great human weakness to generalize from exceptions. As scientists, we know that the best solution to a specific problem is a specific solution. We should attempt to prescribe regulatory procedures that accomplish our objectives without interfering with research. In the two specific instances cited above, the cancer patient and the clinical cancer investigator must insist that the regulatory procedures are altered so that they encourage rather than interfere with progress in developing better treatments for cancer.

### Health Care Delivery

A second major obstacle to clinical research is the current clamor for “primary health care.” We are led to believe that the major problem for our health care system is applying our current knowledge to our entire population. The clinical scientist is viewed as a super specialist, completely isolated from the problems of illness and health care. He resides in his ivory tower, and he deals with only one part of the patient rather than with the entire clinical problem. He is a cold-blooded technocrat without emotion or humanity. There is almost certainly some basis for concern about this problem. There almost certainly is inadequate attention to primary health care, and there almost certainly are clinical investigators who have such characteristics.

As a guide for clinical investigators, I have formulated Freireich’s Law Number Six (Fig. 3). If there is documented progress in terms of improved results from the discovery of new treatment, then it should follow that the patients participating in such research are themselves the beneficiary of those new advances, in contrast to those patients not participating in clinical research. Particularly in the hands of clinical scientists with proven records of achievement, this law is most likely to be fulfilled. There is, of course, a scientific risk of less favorable outcome from any new procedure. However, all of our clinical research procedures are set up to minimize the possibility that changes in therapy will give poor results. The second part of Freireich’s Law Number Six, or the alternate form, states more clearly that any clinical research that does not offer the patient the best or better patient care than could be accomplished outside of the clinical research setting does not qualify for good clinical research. If we can bring together in the minds of the public, of our patients, of the physicians who refer patients to participate in our research programs, and of the administrative people and legislative people in the community that Freireich’s Law Number Six is in fact the guiding principle of clinical investigators, then I think the apparent swing in emphasis away from clinical investigative care toward health care delivery can be modulated. We can allocate resources consistent with the maximum rate of progress in clinical research while still leaving sufficient resources to continuously upgrade the quality of care and to shorten the reduction to practice time for new treatments of proven value.

### Priorities for Investigation

Here I am concerned about generalizations relating to what problems need to be investigated. Examples include the statement and the decision that we should invest all of our resources in working on lung cancer, because it is the most common cancer, and, at the other extreme, that we should invest more of our resources on the study of human acute leukemia, because it is a prototype of other malignant diseases and will lead to a more rapid solution of the cancer problem. Then there is the appeal to the type of patients being affected. For instance, we are told that it is more important to investigate cancer in children, because that will have a greater impact on health than...
cancer in the elderly. Perhaps we are admonished to spend our national resources on cancer prevention rather than treatment or even on basic research as opposed to clinical research. How can such priorities be established? All of these are very complex matters to which I have no clairvoyance; therefore, I cannot make a final judgement, but in choosing priorities, it should be helpful to be guided by Freireich’s Law Number Five (Fig. 4).

This is an adaptation of a frequently quoted admonition that the physician should, as his first principle, “do no harm.” That has always seemed to me to be a particularly offensive admonition for a physician, because to fulfill that requires no action at all. Certainly, any layperson is fully qualified to “do no harm.” The physician’s admonition must clearly be “do what can possibly be done and, perhaps more importantly, do that which is necessary.” We cannot turn our backs on any part of the cancer problem. We must investigate problems in the clinic as they represent themselves. We are required to care for the elderly and for the young, for acute leukemia and for lung cancer. In setting priorities, there is no question that rare resources must be allocated; however, overgeneralization in this area is dangerous for all, but fatal for the clinical scientist. We live in an era of public acceptance of concepts such as euthanasia, abortion on demand, and mandatory sterilization. Frequently, the physician is cast in the role of prolonging suffering or inflicting agony on his patients. It is clear to me that the physician needs no defense. His primary mission is the relief of suffering and the prolongation of life. For the clinical scientist, it is this component of his professional activities, that is, the physician component, that should take precedence. He cannot fail to incorporate the principles of Freireich’s Law Number Six, that is, the best possible care, and he cannot turn his back on clinical problems that require attention and investigation. The clinical scientist must be in the vanguard of physicians, emphasizing the prospects for a continuously improving outlook and for the achievements of clinical research.

Statistical Tyranny

Science in general and medicine in particular are now going through an era in which the great advances and great power of statistics have begun to acquire an air of tyranny. Clinical investigators have begun to think in statistical terms. This has led to highly developed clinical trial procedures, which are now widely adopted. In this section, I would like to examine the components of the clinical trial technique and evaluate each individually and analyze its contribution to discovering new treatments.

Every medical student and physician is thoroughly familiar with the important ingredients of the “controlled trial” (Fig. 5). Each ingredient has a strong intellectual basis for its importance, and results obtained in such clinical trials tend to gain wide acceptance rapidly in the medical community. I was personally an active participant in one of the first cooperative clinical trials conducted by these principles, performed at the NIH under the leadership of Drs. Frei, Zubrod, Holland, and Pinkel in the mid-1950s. In addition, I was the senior author on what is to my knowledge the first placebo-controlled, randomized concurrent clinical trial in cancer therapy. Therefore, I think that I have a deep appreciation of the proposed virtues of such trials. Unfortunately, over the last 20 years, I have also become aware of very serious limitations. These limitations are so serious that I would propose at the present time that there are few, if any, indications for using the classical clinical trial strategy to evaluate and discover new treatments for cancer. I wish to review these components briefly to clarify not only their limitations but also the prospects for improving them.

The first item listed in Fig. 5, prospective, was coined to reflect the fact that research is planned in advance and conducted according to a planned strategy. However, in many people’s minds, it has come to mean something quite different: that is, that knowledge is accumulated in prospect rather than analyzed after the events have occurred. That concept, of course, is ridiculous. There is no prospective knowledge. The only understanding we have from a prospective study is gained from a retrospective analysis of the data. If the prospective study is ever to have any usefulness to anyone, it will be useful only as historical data, i.e., the historical controls with which to compare observations of the future. The prospective study is as likely to give misleading results as any retrospective study. What is crucial is not the prospective or the retrospective aspect, but the quality of the research being conducted. Whereas planning is always helpful in research, in the past, retrospective or historical controls have always provided the only tools for making predictions about the future.

The second item listed in Fig. 5 is randomization. It is intellectually appealing to introduce a technique that will control all variables that are not identified before a study. The technique
of assigning treatment at random has the potential to accomplish that objective. Unfortunately, the allocation of two treatments at random in a comparative study assures comparability of treatment groups only after the study of large numbers of patients. Frequently, small numbers of patients are evaluated in clinical trials. Because of the enormous variability in patients with disease, investigators frequently leave the distribution of known variables to the randomization process. This virtually guarantees that the groups will not be comparable. It should be clear that if 20 or more variables affect response, which is common in clinical studies, then at the 5% level of significance, one would expect a statistically significant difference between the two groups for 1 of the 20. Although on the average the two treatment groups will be comparable, major differences between treatment and control groups will regularly occur. In sequential studies, particularly if one has the advantage of large numbers of historical patient data from which to choose appropriate controls, it is possible to select patient results that give a more comparable control group than can any randomized study.

To overcome the tendency toward noncomparable groups in a randomized study, randomization is frequently supplemented by stratification. The result is that as the number of categories for which patients are stratified increases, the effectiveness of randomization is progressively diminished, because the number of patients in each stratified group will be reduced. For instance, in patients with breast cancer, one can identify at least 10 pretreatment variables known to affect the outcome of treatment. In acute leukemia, more than 15 can be identified. It is clear that to get significant numbers in each of the stratification groups will require clinical studies involving thousands of patients, which are unattainable, even with the most ambitious randomized controlled study.

The next component, the use of a blind, was introduced to allow investigators to evaluate data that required subjective evaluations. For instance, in studies affecting mood, pain, or other subjective elements, placebo controls have proven useful. The extrapolation to cancer therapy studies has, however, created ridiculous situations. We usually use objective measurements of tumor regression, measures of peripheral blood values estimated by instruments, and characteristics such as survival from diagnosis, performance data, and other clearly objective criteria that cannot be influenced by observer bias, unless the observer is frankly dishonest, in which case, the blind will not help at all. A patient with a desperate illness, such as cancer, consults a physician for advice. The physician may offer the patient the fact that there is no treatment of proven value or that the treatment of proven value is of sufficiently limited effectiveness that clinical research is indicated. After explaining in great detail what the side effects and risks of such new treatment are, he must finally ask the patient to consent to receive either the new treatment or the conventional treatment. In such a setting, it is hard for me to believe that many fully informed patients would in fact consent to such treatment. It seems clear to me that the physician who recommends to his patient that he receive a new or developmental therapy should be convinced before he makes that recommendation that, based on all the available knowledge, the probability is greater for benefit than risk. In the absence of that conviction, I do not honestly believe that the physician should offer the patient the extra risk of developmental therapy. When the physician’s judgement is that the benefit:risk ratio favors the developmental therapy, then to ask the patient to consent to receiving therapy that in that physician’s judgement is inferior would not be honest. Many patients and many legal and other community organizations are now beginning to recognize the serious limitation of blind studies for clinical investigation, and it is my unambiguous view that such studies should be done rarely, if ever.

Perhaps the most limiting concept in the clinical trial is the need for control groups that are studied concurrently and comparatively. It should be emphasized that there are very significant questions to be asked that do not involve comparison but rather estimation of the degree of effectiveness of the treatment.

In a comparative trial, when significant differences are declared, one is frequently left with a poor estimate of the magnitude of the effectiveness of the new treatment. Fig. 6 illustrates a study of the weight of a sample of material, which, on comparison to the known weight, is clearly heavier. The difference is large, so the significance is easy to establish, but the actual magnitude of the difference is unknown. In the second panel, the same sample is measured on a scale that gives a precise weight. The scale has been previously calibrated based on the historical data, that is, the use of standard weights, which are converted by the spring or the counterbalance method to a scale that gives predictable estimates of the weight. In clinical research, it is probably more important to know what the probability of response is for a given treatment than to know that a given treatment is either better or worse than another therapy.

The requirement for concurrence imposes serious restrictions. The individual performing on the clarinet can make beautiful music and can, in sequence, put down the clarinet and play the saxophone quite well. The individual who is attempting to play multiple instruments at one time frequently ends up with dissonance. Likewise, in clinical research, in the absence of clear evidence or a reason to suspect that the disease is changing with the passage of time, there is every reason to have confidence in the historical control and to make accurate estimates of the effectiveness of a new treatment subsequently compared to the historical control.

Few recognize the tremendous cost of asking estimation questions in comparative form. To illustrate this, I have designed a clinical trial (Table 2) which assumes that we know from a very large number of patients, perhaps thousands, that the response rate for conventional treatment was 20%. We are evaluating a new treatment and decide in advance that an increase of 100% in response rate, that is, a 40% response rate, would be a significant advance in our therapy. I have assumed that this new treatment has a true overall response rate of 40%, and therefore, it would definitely be worthwhile to declare this as a superior treatment by our own criteria. In this specific example, we have set the numbers so that the best of all possible worlds is in fact realized. Thus, after every increment of 10 patients added to the study, the observed response rate will be exactly correct, that is, we will always have exactly a 20% response rate for the conventional treatment and a 40% response rate in the experimental therapy group. Of course, this situation would never be realized in actual practice. There will be estimation errors in both groups. I will return to that point in a minute. Ignoring the measurement error, after we have entered
20 consecutive patients on a sequential study and compared the exact response rate of 40%, that is, 8 responses in 20 patients, to the already established value of 20% from our historical control, we would have a $\chi^2$ value of 3.83 and a probability, that is, a $P$ value, that this was significantly different from the historical control at the 5% level of significance. If we did the same study in a comparative fashion, we would allocate conventional treatment to half, or 10, of the patients, and half, only 10, would receive the new treatment. Even though we observed the exact response rate of 20 and 40%, we would now have a $\chi^2$ value for comparison of only 1.24, which gives a $P$ of 0.63. This would have resulted in our declaring erroneously that there was no difference between these two treatments. As you can see from the table, to reach a level of significance of approximately 5%, we would have to study 100 patients or five times as many patients in the comparative trial as were necessary in the sequential trial.

Because, in a comparative trial, there may be as good a possibility of an inferior outcome as a superior outcome, we may do the patient a favor by giving him an equal chance of receiving the control treatment. One might therefore say that a comparison of a sequential and a comparative study in this fashion would not be fair, because it only considers the favorable outcome. However, even if you were to be willing to study twice the number of patients, that is, 40 patients, you would now have 20 patients on the experimental therapy, the same number that was observed in the sequential trial, and when one computes a $\chi^2$ comparing the 20% response to the 40% response in two 20-patient groups, you still have a $P$ of only 0.30, and you still would have rejected this treatment as not significantly better. The reason for this is that you have given up the knowledge of the historical response rate and have again selected to make an independent estimate of the historical data with a small sample. Few people consider that in a concurrent clinical trial, one is just as likely to get a poor estimate of the control treatment as you are of the experimental treatment. Thus, the comparative trial can be misleading not only because of the necessity to commit larger numbers of patients to the study, but because it increases the possibility of reaching the wrong conclusion, because measurements are made of both the historical control or conventional treatment and the new treatment simultaneously.

Table 2  Significance levels by sample size for study comparing 20% historical response rate to 40% new treatment response rate

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Randomized study</th>
<th>Historical control study</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.98</td>
<td>0.24</td>
</tr>
<tr>
<td>20</td>
<td>0.63</td>
<td>0.05</td>
</tr>
<tr>
<td>30</td>
<td>0.43</td>
<td>0.01</td>
</tr>
<tr>
<td>40</td>
<td>0.30</td>
<td>0.003</td>
</tr>
<tr>
<td>50</td>
<td>0.22</td>
<td>0.0008</td>
</tr>
<tr>
<td>80</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

Perhaps the most important impact of statistics on the clinical investigator's judgment is what has been called the statistical overkill, that is, such a strong reliance on statistics that the clinical and biological facts surrounding those observations are either overlooked or ignored. For instance, in comparing two treatments, an investigator may find a statistically significant difference in response rate, which is of absolutely no biological significance. If a sufficiently large number of patients are treated without attention to stratification for prognostic variables, pooling patients from multiple centers and even from multiple studies, sufficient numbers can be obtained such that differences of small magnitude are statistically significant. Other factors could be much more important in deciding which treatment the patient should receive, factors such as cost, unpleasantness, toxicity, availability, and so forth. In contrast, at the other extreme, are observations of great biological significance that have no statistical significance. There are an unlimited number of examples of circumstances in which historical data reveal very infrequent response with conventional treatment. A high proportion of responses in a small number of patients may have no statistical significance but may be of enormous biological significance and therefore well worth reporting. Have we, as clinical investigators, overlooked the fantastically significant biological contributions that have been made to our knowledge of medicine from observations in single patients, the studies of the physiology of the stomach in the wounded soldier, the understanding and detection of metabolic and endocrine disturbances, the contributions to biological understanding of health and disease that derive from the study of a single individual? We still have in our medical literature an important place for the "case report," experiments of nature that point the way toward dramatically new, significant, important observations. Statistics cannot prove anything or disprove anything. Proof depends not only on statistical significance, but on a whole range of biological observations, including, most importantly, its value for predicting the future, that is, the reproducibility of that knowledge. Statistics assist the investigator in helping to understand the results of his investigation. It still requires a capable, intelligent clinical investigator to recognize and identify important differences.

My friend and colleague Dr. Edmund Gehan has taught me what I have adapted as Freireich's Law Number Four (Fig. 7). Notice that it does not say that the best therapeutic research gives the highest $P$ or the most statistical significance but the best overall treatment results. The statistician must assist the
clinical investigator in his analysis and quantitation of his data. The statistician must share with the clinical investigator the same objectives of identifying the best clinical results for the highest proportion of patients. Research must be judged by the significance of the discovery, how significantly it differs from the past, and what the overall benefit of such treatments is to the individuals who have been treated and will be treated in the future.

**Academic Requirements**

Next I want to examine the academic requirements for clinical investigation. To conduct clinical research, most of us turn to federal and community sources for support. In the process of peer review, not only for funding, but for publication of our scientific data and for our status in the scientific community, we find a whole new set of less objective and more emotional or subjective criteria that characterize the clinical trial.

We have seen our colleagues and critics characterize the clinical trial as "scientific" and other types of clinical investigation as "unscientific" (Table 3). Many academic review bodies are prepared to reject information that lacks the imprimatur of the controlled clinical trial. Authors have evaluated therapeutic research based entirely on whether or not randomization was used. In fact, some have confined the term "controlled" to studies that have the characteristics of prospective, randomized, and concurrent studies and frequently refer to other types of investigations as "uncontrolled." This is a particularly malignant allegation, because there are many circumstances in which historical controls, matched controls, and other types of controls are in fact superior to prospective concurrent controls. We have witnessed many examples of national multi-institutional cooperative, collaborative studies of disease so rare that the anticipation is that over 4 and 5 years, the number of patients that will be studied will be less than could possibly detect a major difference between a new treatment and a conventional treatment. Yet, in slavish adherence to the ritual of the clinical trial, investigators have elected to conduct their trials and publish data suggesting that even though the numbers were small and the data were not statistically significant, the results should be observed. How much better it would have been for them to have studied their patients in sequence to get an excellent estimate of the effectiveness of the new treatment.

The phrase "controlled clinical trial" does not belong to "THE clinical trial." Controls are essential in clinical research. The problem is the technique for selecting controls. Appropriate controls can be selected by a wide variety of techniques and are improving each day. In fact, the new strategy of a statistical, stepwise, forward regression analysis of multiple pretreatment variables to compute the distribution of response probabilities in treated populations and in control populations has given much better comparability between treatment and control groups. Let us admit that there are many ways to conduct scientific and "controlled" studies. We have heard many times at meetings, and we have heard many times today, investigators report preliminary results on groups of patients. They frequently conclude by saying "Of course, this needs to be confirmed in a randomized controlled trial." The presumption that we must continuously doubt the data that we have unless the formal clinical trial strategy is applied is so deeply ingrained that even investigators who realize that it is preferable to do sequential trials still must concede in their final analysis that controlled trials are indicated. In reality, once a sequential trial demonstrates that the new treatment has a significant probability of being substantially better than the historical treatment, the prospective comparative trial can actually be significantly counterproductive and, in fact, contraindicated.

The next three items on Table 3 are closely related. The penultimate in the clinical trial has now become the national study group. There are many of these national studies, generally NIH sponsored, that represent cooperative groups of usually university-based physicians who use highly designed prospective randomized clinical trials that are usually supervised and operated by a study office, generally administered by a statistician or by a group of statisticians. Publications from such a group are frequently anonymous or contain a footnote indicating that there was either an evaluation or a writing committee and that frequently lists the participating institutions. The conclusions derived from such studies are presumed to have permanent, unambiguous, conclusive capacity to establish the truth. This is frequently the case, even though previous and subsequent clinical experience is in complete disagreement with the results of such studies. One of the earliest and perhaps the most famous of such collaborative studies was the UDGP study. This study was initiated in 1961, and the results were published more than a decade ago. The results of the study led to the conclusion that oral hypoglycemic agents were harmful rather than helpful, much to the astonishment of the entire biomedical community. There has probably been more published controversy about the data contained in that study than in any other previous single study in the history of clinical investigation. In addition to the publications, pro and con, a panel of statisticians was appointed to provide final judgement on the study, based on a complete reanalysis of not only the data but of all the objections and criticisms raised in medical literature; hence, the ultimate in statistical tyranny. We found that many investigators were prepared to accept that a committee of statisticians was best qualified to make a judgement on an unambiguously medical therapy matter. The committee determined that the study con-

**Table 3**

<table>
<thead>
<tr>
<th>“THE” Clinical trial</th>
<th>Clinical investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Scientific”</td>
<td>“Unscientific”</td>
</tr>
<tr>
<td>“Controlled”</td>
<td>“Uncontrolled”</td>
</tr>
<tr>
<td>National study (NIH sponsored)</td>
<td>Single institution</td>
</tr>
<tr>
<td>Cooperative group (anonymous)</td>
<td>Principal investigator</td>
</tr>
<tr>
<td>Statistical tyranny</td>
<td>Statistician collaborator</td>
</tr>
</tbody>
</table>
clusions were correct and that the controversy should be entirely settled, but unfortunately, it was not.

One individual who has made significant contributions to the biostatistical literature, Dr. Feinstein, a physician and a statistician, has written a series of papers on the subject of that study. This quote is taken from one of his recent publications.

"The main lesson to be learned from this study had nothing to do with the virtues or evils of the hypoglycemic agents. The main lesson is the simple truism that a clinical trial of therapy must be designed as a CLINICAL trial of therapy."

He points out in that article that the investigators who participated in the study were university professors who would not have taught their students or granted board certification to any physician who practiced medicine in the fashion prescribed in the study.

Dr. Lasagna, who has contributed more to drug evaluation than probably any other investigator, has recently written a satirical article on drug development. The following quote is taken from that article. "Some critics have said that the UDGP study showed only that if you give oral hypoglycemic agents to the wrong people, in the wrong way, that the agents will be toxic. To that I say, picky, picky, picky."

Indeed! A poorly designed clinical trial cannot be salvaged by statistics, nor can a well-designed study be invalidated by statistics.

Despite this voluminous literature on the clinical trial of the UDGP study group, it is astonishing that there has been little effort to either reproduce or confirm those results, which is the scientific strategy that has proven to be the most reliable and dependable for resolving questions, rather than prolonged, heated emotional debate. The national cooperative groups' statistically valid studies have many limitations. One referred to by many authors is the fact that there is frequently greater variability in results between cooperating institutions than there is between the treatments being compared. Isn't it likely that there is, in fact, a qualitative and quantitative difference between the excellence of care, of adherence to protocols, and of knowledge and experience between institutions, which is as much responsible for the differences observed as are the treatments being compared?

I am not convinced that the best quality of scientific investigative achievement derives from the deliberation of committees. How many times have we attended meetings of committees with carefully thought out, entirely original ideas, only to find that others had entirely different approaches to the problem? Then the committee action reaches a consensus that is somewhere between sharply divergent objectives. Such compromises frequently frustrate the good objectives of both authors. I recall the story of a commercial organization that was about to introduce a remote control device for picking up things. They did a market survey and determined that there was one very large group of potential consumers who wanted to have a short arm and another very large group of consumers who wanted to have a large arm. The company came to the compromise decision to manufacture one of intermediate length. The result, of course, was that none were sold, because it was not appropriate to either group. Similarly the consensus-type research that goes on in committees is frequently stifling to the creative individual, who has the capacity to design highly innovative studies.

Another major factor in academic review is what I call "the good is bad, bad is good principle." I have been involved in two major clinical research projects where this principle has been invoked as one of the crucial criteria for disproving both funding and publication of this research. The first was in the area of granulocyte replacement transfusions, and the second was in the use of reverse isolation procedures to diminish infectious complications in patients undergoing chemotherapy. In both instances, we had found by treating patients in sequence that the eventual outcome by objective criteria such as survival, frequency of complete remission, frequency of infections, and so forth was significantly and substantially better than the historical control data. The reason that concurrent randomized studies were not undertaken was that in both instances, the technology was expensive, and the number of patients that could be studied was limited. Therefore, it was clear that if we limited our studies to concurrent studies, we would be unable to detect significant differences for many years. Moreover, the technology was being improved by the results being obtained during the research, and we therefore felt that it was not sufficiently well developed to justify a prospective, highly designed, randomized, comparative study.

Figure 8 shows diagrammatically how the principle is used. In the upper panel is a report of a historically controlled study. A group of patients were treated in sequence, and the results, using objective and quantitative criteria of response, were compared to our historical data. The outcome was substantially and significantly better than had been reported in the past by many authors, including our own data. As I already pointed out, the academic judgement of such a study was that the research was bad. Although it gave good results, the main weakness was the alleged poor experimental design, the failure to use "the clinical trial" technology. Over the next 10 years, there were several publications and investigations in which other investigators had conducted classical clinical trials, randomly allocating half of their patients to the conventional treatment and the other half to the proposed new treatment. In both instances, these studies unambiguously confirmed the positive results already published. They did, however, have a bad result, that is, half of their patients had unfavorable results. In fact, in one study, the criterion for failure was death. By ordinary, intelligent criteria of clinical judgement, one would have to say that this was a bad clinical result, at least for half of the patients, yet the academic judgement was that this was good research, resulting in adequate funding and prompt publication. Such unjustified confidence in the intellectually satisfying principles of the clinical trial cannot take precedence over our critical and intelligent judgement of scientific results. "The good is bad, bad is good principle" should be violated by all clinical investigators.

Freireich's Law Number Three (Fig. 9) suggests that it cannot be necessary to have a bad result before we can be convinced of the good results. It is my view that we will have to be satisfied with the knowledge that the results are better than expected and be satisfied with confirming this in subsequent studies and not be forced to experiment on patients by offering treatments in which the expectation is for a poorer outcome. Academic questions can be addressed very effectively in the laboratory, but not in man.
THE "GOOD IS BAD, BAD IS GOOD" PRINCIPLE:

![Diagram of Result Evaluation]

False Positive–False Negative Choices

The next area I wish to discuss is the decision made by the clinical investigator as to the type of risk that he will take when undertaking a clinical study. For any investigation, regardless of its size, there is a probability of declaring a treatment effective at predetermined levels of significance when, in truth, it is not effective, the so-called false positive error. At the same time, there is a risk of declaring a treatment inactive when, in fact, it is active by our criteria, the so-called false negative. Statisticians refer to these as Type 1 and Type 2 risks, but I think the words are more descriptive. How should the clinical investigator make such choices, and what criteria should be used?

The classical approach to this problem by statisticians is the use of the null hypothesis. The investigator hypothesizes that no difference exists, and the test is made to reject the hypothesis. If the data force the rejection of the null hypothesis, the investigator should maintain a positive inquisitive attitude. His greatest fear should be a false negative result, and his efforts will be aimed at protecting against that risk. A false positive result will lead quickly to an effort to confirm. These efforts will be undertaken both by the same investigator in larger groups of patients, by other investigators in their own clinics, and ultimately by interinstitutional cooperative efforts. If, in fact, the result is not important, subsequent study will soon reveal that this is not a positive result. Contrast this to the false negative result. Whenever an investigator declares that a treatment is ineffective, there will likely be no confirming studies undertaken. I doubt that any clinical investigator has rushed to undertake a clinical trial with drugs that have previously been evaluated and found to be ineffective. The false negative result is therefore usually fatal for the development of the treatment. It is still true that in the processes of science, it is only the positive information that allows us to make predictions. Careful analysis of historical data and thorough understanding of its implications allow the investigator to make predictions of the future that can subsequently be confirmed by observation and by testing. There are no methods for proving that something is ineffective and inactive. There is no intellectual way to prove the negative. We can fail to discover the positive laws and information that will predict for the future, but we cannot disprove claims. For those who do not intuitively accept that statement, I need only to refer you to the tremendous turmoil that science and scientists go through when faced with claims made by quacks, faith healers, and other people of this kind, who claim activity but have no scientific data to substantiate it. Many clinical trials have been undertaken to demonstrate the lack of usefulness of a drug. When such a trial is undertaken, the risk of a false positive exists. Of course, that aggravates the problem. Perhaps more important is the fact that no matter how negative the study, the claim will still stand, because one can always contend that the treatment was not given precisely right or at the right time of day and so forth. What I am trying to say is that the clinical investigator should maintain a positive inquisitive attitude. His greatest fear should be a false negative result, and his efforts should be aimed at protecting against that risk. A false positive result is no problem. I remember in the early days of chemotherapy, my medical colleagues would inquire about the “poison of the year,” the month, or the week. The implication was that new treatments were voguish and that we were always falsely optimistic about our results. What these cynical individuals failed to notice was that the drug of the month was almost invariably replaced by better drugs, and that in the last 20 years, there have been dramatic changes in the effectiveness of our treatments for metastatic cancer, for adjuvant treatment and for all phases of cancer.

Another concept that tends to emphasize a pessimistic view of clinical research is the two-sided testing of the hypothesis. If one is investigating a new treatment, the only intellectual basis for undertaking such research is to seek a superior form of treatment. Yet we frequently set up our clinical trials such that we have the same power of discovering that this is either better or worse than conventional treatment. It seems ludicrous to me to attempt to establish that treatment is inferior. Frequently, in comparative studies, after a number of patients have been studied, it is evident that the probability that the new treatment is significantly better is very low; at that point, it would seem to me to be patently clear that further investigation is not needed. Perhaps more important is that the statistics of the one-sided test allow us to detect significant positive differences more rapidly, without requiring that we be equally sensitive to the negative results. In my view, particularly when seeking better treatments in comparison to conventional treatment, only one-sided testing is indicated.
In the conventional comparative clinical trial, few investigators realize that this experimental plan has a high propensity to give false negative results. One reason is that with rigid protocols and with pooling data across stratification and prognostic variables, differences between average patients are often minor, whereas significant improvement is observed in a smaller fraction of patients. I have frequently alluded to “median disease” only to emphasize that many statistical and clinical research concepts revolve around the average patient, yet many of our treatments are beneficial for 20 or 30% of our patients and are still clearly effective and justifiable. We always need to be alert to identify the characteristics of those patients that do respond favorably to treatment and to carefully select patients for treatment in accord with those characteristics. Of course, ultimately, one would like to have treatment choices tailored to the patient’s own characteristics and his own pharmacological handling of the drugs and to have a prescription for him that offers the greatest probability of response. At the present state of our knowledge, this is not always possible, but on the other hand, the propensity to study only large numbers of patients and to prescribe only for average patient groups has a high propensity of declaring treatment to be inactive, which in fact, has value.

I have always been impressed by the fantastically high regard that clinical scientists and statisticians hold for the 5% level of significance. When I searched for the origin of this, I discovered that the famous Dr. Fisher had casually mentioned in his book that it seemed reasonable to choose levels of significance of approximately 5%. It seems clear to me that a drug that has a 90% probability of being better would be my personal treatment choice. Again, our worshipping of the 5% level of significance frequently leads us to make false negative errors; it leads us to reject treatments simply because the number of patients evaluated is too small, and $P$ exceeds 5%. I have actually heard an investigator say when treatment A differed from treatment B, with a $P$ value of only 0.06, that the treatment was not significantly better and that we should discard the treatment.

In approaching the choice between false positive and false negative risks, Freireich’s Law Number Two has proven useful (Fig. 10). Most of us expend great efforts preparing for failure. At a very young age, we buy life insurance to prepare for death and automobile insurance to prepare for catastrophes. We even become used to testing new drugs and finding them inactive. But few of us are prepared for the favorable outcome. Are we prepared to introduce a new drug tomorrow and discover that all of our patients will have their cancer permanently and irreversibly cured? I am not sure that we are. I recall the kinds of crises that followed the dramatic breakthrough of the polio vaccine because we could not produce sufficient vaccine. If we are ready and open-minded and observe in our clinical study those results that could be favorable in our patients and make every effort to avoid false negative results by interpreting even the most minor changes as potentially beneficial, we will have the greatest prospect of making progress. If the clinical investigator is not optimistic in his choice of new treatments for his patients, who in the health care system will be? We have to offer this optimism to our patients so that they also feel that the drugs to which they are being exposed and the treatments that they are receiving do have prospects for dramatic changes in outcome for the better. There have been dramatic breakthroughs in the clinical investigations of cancer, and virtually all of these have been initiated by optimistic individuals in sequential series of patients making quantitative observations, a strategy that has been effective in the past and will be effective in the future.

**Ethics**

The subject that I reserved for last is the ethical aspects of clinical research. In one sense, this is the easiest subject to discuss, because we are physicians and scientists. We recognize our ethical commitments to our patients, to our institutions, to our community, and even to the world of man. It would be almost boring for me to pretend to have anything significant to say about the matter, so I will resist the temptation and simply offer the formulation of Freireich’s Law Number One (Fig. 11).

This is the most important of the seven Freireich Laws for clinical investigators, and it is therefore appropriately known as the Clinical Investigator’s Creed. I have heard many scientists say that in treating our patients, we have a significant commitment to the people who will develop cancer in the future, to our scientific colleagues, and to other physicians to be certain of the outcomes. Therefore, certain types of experimental plans are justifiable. In my view, this is the first step on a slippery slope toward experimenting on people that ends in violating all of the other six Freireich Laws. The first consideration of a clinical investigator must always be the welfare of that person. Other considerations, although important, must always be secondary. That ethical position, it seems to me, is inescapable. It is my firm belief that every patient must have absolute confidence that his physician has a primary interest in his medical problem and that the medical problems of others are of secondary importance.

This ends my discussion of the seven major obstacles to innovative, productive, clinical research. I realize that much of this presentation has been critical, however, I have attempted, in the form of Freireich’s Laws, to formulate guideposts that would assist in resolving and overcoming these obstacles.

I am sure that many of you will judge my presentation as regressive—it might seem as if I were supporting a retreat from the technology and scientific methodology that has been so important to our current stage of development and knowledge in clinical care. Such regression might be visualized as the swinging of a pendulum between clinical trial and clinical experience. Some of you may misinterpret my talk to indicate that the pendulum has swung too far toward the clinical trial and objective, quantitative, coolly scientific technology.

Perhaps I have indicated that it is time for us to move back in the direction of more observational, individualized medical
FREIREICH’S LAW # 1
(Clinical Investigator’s Creed)

The primary beneficiary of clinical research is the patient participating in that research.

Fig. 11

treatment. However, I do not wish my talk to be misinterpreted as an anti-intellectual call for a return to the anecdotal type of medicine that we were taught in the 1940s. As you all know, I am a strong advocate of scientific methodology, of research, and of rigor and discipline in our work. I thought long and hard about how I could clarify the position that I have taken before you today.

A better analogy than the pendulum is the tacking of a sailboat as it pursues the direction that the captain decided to go, with the destination of our community set on the control of cancer. It is often not possible to go directly to the target because of the direction of the wind. It is necessary to swing the boat toward progressively more scientific, more ordered, more planned research. I think we have done that, and we have gone far enough in that direction. If we continue to go further, we will veer off our course. I think the time has come to change directions, to swing the boat 90 degrees back toward the type of clinical research that is more observational, and I propose that such a change will keep us relentlessly on target toward our goal of cancer control.

Who took the clinical out of clinical research? Many individuals and institutions suggest that clinical research be practiced as laboratory research. There are now and always will be many individuals and forces that detract the clinical scientist’s attention from his commitment to the clinical in his research to the relationship between physician and patient.

It is the clinical investigator’s responsibility to assure that the important characteristics of research on patients receive first priority in his approach to scientific problems. Like other scientists, the clinical investigator must have extensive knowledge of the subject being investigated. Before any developmental therapy is evaluated, the physician must have a clear expectation of what the outcome would be with conventional treatment, what the prognosis for such a patient is, and what the appropriate risk:benefit ratios are that should be considered. The clinical investigator must be experienced in treatment and observation of the disease being investigated, preferably by direct clinical experience in the one-to-one situation between physician and patient. The clinical investigator must be a skillful physician. It is certainly true that the levels of skill of the investigator will to a large degree determine the quality of his research, rather than the research techniques. Finally, the clinical investigator must have a past record of achievement and of innovativeness.

It is this physician–knowledgeable, experienced, skillful, innovative, possessed of excellent clinical judgement and humanity, and whose constant emphasis is on the physician, that is, the clinical aspects of research—who has the characteristics to which all of us should aspire.

Reference


E J Freireich