11 Extrapolation of trial results

As the Renaissance in Europe was drawing to a close, some gamblers were caught up in the new spirit of inquiry. They became curious about the pattern of numbers that kept turning up in the fall of dice. Unable to answer their own questions about these seemingly fateful happenings, they approached some of the leading scientists of the day for help. A group of Italian gamblers met with an interested reception when they consulted Galileo. The busy astronomer-physicist found the gaming problems worthy of study, and he went on to write a treatise on die-throwing sometime between 1613 and 1623.

A short time later (*ca* 1654), history repeated itself in France. A famous gambler, the Chevalier de Méré, was perplexed about 'the problem of points in the die game'. He brought the questions to his friend, Blaise Pascal, the mathematical genius. Subsequently, Pascal and his correspondent, Pierre de Fermat, in a famous exchange of letters, developed a solution to the problem and they generalized the result in the form of tables that indicated how stakes should be divided in games between two players if they decide to quit at a moment when neither has definitely won. These 'beautiful ideas about games of chance' were admired by other mathematicians and led directly to the development of the theory of probability.

PRACTICAL ACTION IN THE FACE OF UNCERTAINTY

I find it interesting that the new theory arose in response to the needs of those who risked their fortunes daily on actions taken in the face of uncertainty. And considering the number of parallels between gambling and medicine that I have drawn on, I am not surprised that the regular workings of the laws of chance had been anticipated in the mid 1500s by a physician, Giralomo Cardano.

State of nature

The theory of probability may be defined as the logic of degrees of belief concerning the occurrence of uncertain events. But what do we mean by

Giralomo Cardano (1501–76: physician, mathematician, philosopher, and gambler) was the first to calculate a probability correctly (*ca* 1563 or 1564). In an obscure text on games of chance (*Liber de Ludo Aleae*, The Book on Games of Chance, that appeared after his death), he discussed equiprobability, mathematical expectation, reasoning concerning the mean, frequency tables for dice probabilities, additive properties of probabilities, and the question of how many trials are required to give a player an even chance of winning: for example, to throw a given point total in a game of dice.

uncertainty? Recall that there are two kinds of uncertain events. One is due to the kind of randomness exemplified by the fall of honest dice considered by Pascal and Fermat. The other kind of uncertainty arises when the abstract calculation of odds concerning *equi*probable events does not work. The latter are the situations in which we do not know which laws of randomness apply. In statistical terminology the 'state of nature' is unknown. How do we develop a betting strategy for a game involving 'loaded' dice in which one result is favored? As already noted, here we are obliged to undertake a series of observations to estimate the state of nature, but we are immediately confronted by the issue of deciding when we have made 'enough' observations. The judgment depends on the cost of each throw of the dice and the cost of making a wrong decision.

Obviously, we can never be certain about the state of nature no matter how many observations we record concerning the behavior of the dice. For example, it is possible, although highly unlikely, that honest dice will turn up 'seven' in a very high proportion of a hundred consecutive throws. It is also possible, and improbable, that dice weighted to favor 'seven' will turn up 'three' in the majority of one hundred plays.

To evaluate the chances of being led astray by rare results, we turn to the theory of probability. But action-oriented decisions (such as 'our estimate of the state of nature is good enough to place a bet', 'the odds offered are satisfactory', and '$100 is a reasonable wager') cannot be undertaken rationally until we have considered some additional 'facts of life' having to do with values. It is just this kind of situation that faces the physician who must decide on a practical course of action in the best interest of his individual patient.

Disciplined approach to action-oriented decisions

The decision problem under uncertainty is not unique in gambling and in medicine. Recent years have seen the development of practical strategies for balancing risks and benefits when making decisions in states of doubt. These have evolved into the discipline of decision analysis, which has found wide application in fields ranging from military strategy to public health.

On being right or wrong Before discussing the analytic approach to decision making, I should mention that we tend to have unrealistic expectations about the predictions used to make decisions, particularly when the resultant actions produce irreversible effects on our everyday lives. From our perspective as patients, the overriding concern is the correctness of decisions. Alan R. Shapiro of Stanford University, however, points out that we need to distinguish a decision from the evidence on which that decision is based.

In evaluating predictive ability, we must recognize that a final decision is the result of a number of considerations, such as various value assessments and outcomes whose meaning is unique to a particular patient. What is needed in judging predictions, Shapiro notes, is a method for assessing the accuracy of the probabilities asserted, not an all-or-none scoring of decisions as right or wrong.

Frustration and confusion about decisions based on before-the-event predictions come up frequently in clinical medicine. When a surgeon declares before an operation that the mortality risk of the procedure is only 1 per cent and the patient dies on the first postoperative day, the anger of the family and friends is understandable. But it cannot be concluded that the pre-event probability was wrong; without more evidence, we can only say that an improbable event occurred.

Formulation of a decision problem when faced with conditions of uncertainty

1 Specify the viable options available for gathering information, for experimentation, and for action.
2 Specify the events that may possibly occur.
3 Arrange the information that may be acquired in chronologic order, and the choices that may be made with the passage of time.
4 Decide on preferences about the consequences resulting from various courses of action that can be taken (preferences for consequences numerically scaled in terms of utility values).
5 Make a judgment concerning the chances that any particular uncertain event will occur (judgment about uncertainty numerically scaled in terms of probability).
(From Howard Raiffa of Harvard University)

DECISION ANALYSIS

The development of decision theory has provided an analytic guide to the way in which we go about making value decisions in everyday life. The approach lends itself to a logical analysis of decisions in medicine.

First, we need to quantify information and preferences and then proceed in an orderly sequence. Decision analysts advise that a rational choice of treatment requires that we define outcomes and the probabilities of outcomes and that we make some estimate of the values or utilities of the state of health that constitute the outcomes (utility is a numerical measure of worth). Moreover, judgment should be based on the state of each patient as an individual rather than on the disease label.

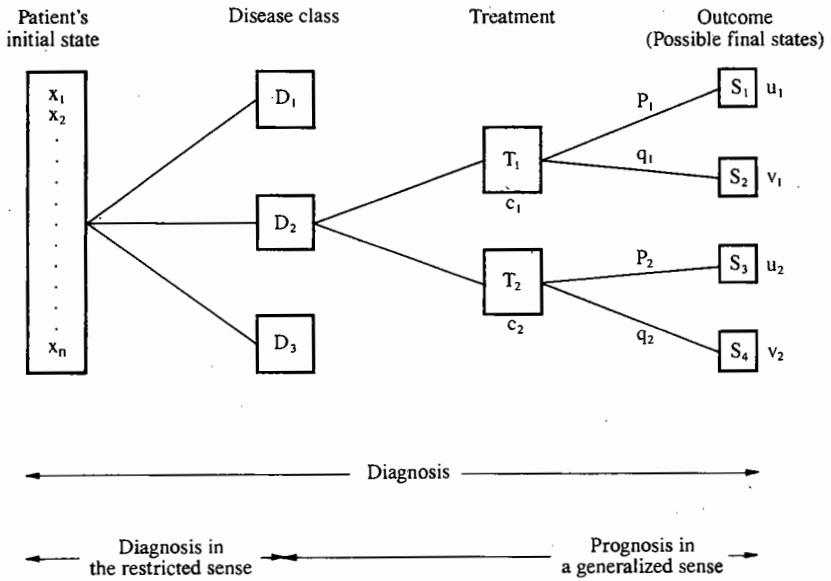
Decision theory in medicine
Definitions of some terms
Decision theory: a theory of decision functions which assumes that it is possible to attach a measure of worth or value to any state of health.
Decision function: action (decided in advance) that will be undertaken as a function of possible outcomes. Ideally, we should like to be able to express possible medical treatments as a decision function whose independent variables are the possible sets of indicants that might occur on any one occasion.
Indicant: any piece of evidence relevant to the probability that a disease is present.
Event: a feature of the state of the patient referring to his state of health.
Event space: the state of a patient is conceived as specifiable, in principle, by a sequence of measurements some of which might be impracticable by existing methods. Each measurement is regarded as a coordinate and the sequence of measurements thus corresponds to a point in multidimensional space called 'event space'.
Utility: the concept of utility (in a sense borrowed from economics where the worth of outcomes is measured numerically; each outcome is given a single number called its utility) is useful in defining objectives in medicine. A doctor attempts to increase each patient's utility, but has no certainty, only a probability of achieving this. Thus it is more accurate to speak of expected utility. The expected utility is the sum of the probabilities of the utilities of all mutually exclusive outcomes. The calculation requires the estimation of the utilities of the states of health of the patient, and the combination of these with the costs of the course of action. Utility is expressed as a mathematical variable or function corresponding to the state of the world for a person. It is assumed that if a person's behaviors and choices satisfy certain compelling postulates the individual will behave as if he/she
- -has subjective probability estimates about the state of the world
- -has associated a real function with possible states of the world
- -wishes to maximize the expected value of the function.

Quasi-utility: when some relevant utilities are too difficult to estimate, it is of value to define some function of probabilities of the more readily estimated utilities whose expectation is estimated instead. Such a function is called a 'quasi-utility'.
Utile: any unit in terms of which utilities could be measured.
Cost: negative contribution to utility; expressed in financial terms or in terms of danger, anxiety or pain to the patient.
Principle of rationality: the maximization of the mathematical expectation of utility, by choice of actions.
Rationality: consistency of choice, and adherence to the principle of rationality.
(Taken from Card and Good)

Diagnostic process

Medical activity, from first examination to active treatment, has been envisioned by W.I. Card and I.J. Good of the University of Glasgow as a sequence of decisions under the general heading of diagnosis. The indicators of disease (called 'indicants') are obtained by questions, laboratory tests, and other procedures at each stage in the overall diagnostic process. A collection of indicants is chosen by the doctor and this selection (among a number of possible sets) decides the next move in the methodic approach.

A medical decision model



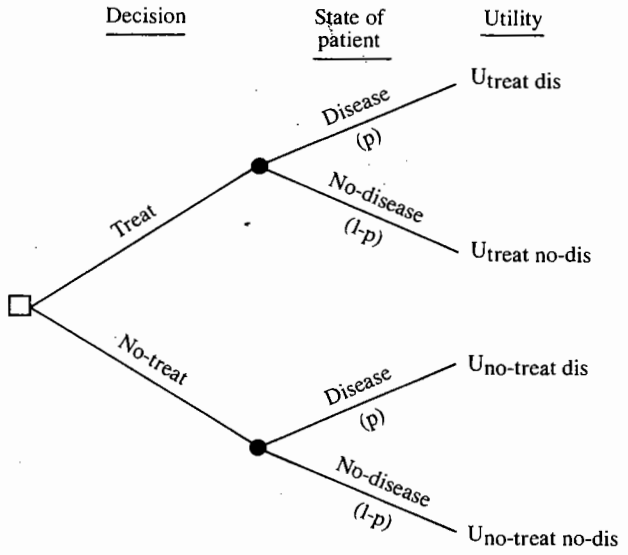
The diagnostic process, Card and Good explain, consists of eliciting a set of indicants from a patient and, on the basis of these, allocating the sufferer to a disease class. Choice of treatment T_1 costs c_1 and with probabilities p_1 and q_1, respectively, produces gains in health utility u_1 and v_1. Similarly, treatment T_2 costs c_2 and with probabilities p_2 and q_2, respectively, produces gains in health utility of u_2 and v_2.

Diagnosis may be regarded as following the branches along a decision tree. The course may be compared to the diagram of a game in which the doctor chooses a move at each stage of the match, and an opponent (the goddess of fortune) chooses a play. The physician's overall search strategy is formed by decisions based on current knowledge, not the whole tree which represents all possible games that could be played.

Gain in utility The rules for the choice of indicants will depend on the initial probabilities of various possible diseases, the expected utilities of their treatments, and on the costs or negative utilities of tests and treatments. A doctor attempts to increase each patient's utility, and the objective is achieved by the choices that are under his control.

In the early stages the gain in utility is indirect: a gain in certainty which, it is hoped, will lead to the best treatment. The estimate of the probability that a specific disease is present may be calculated by Bayes' Theorem

Benefits and costs of treatment choices



In this hypothetical example, the choice to be made by the physician (□) is whether or not to administer positive pressure to a baby with hyaline membrane disease (treat or no-treat). With either choice, chance will 'decide' (●) that the infant either has the disease (with probability p) or does not (probability $1-p$).

Four possible outcomes are represented by the terminal branches of the tree, and each has a certain utility value (e.g. $U_{treat\ dis}$, the utility of administering positive-pressure treatment to patients with hyaline membrane disease). If the four utilities are measured in consistent units (e.g. duration of life*, parental value terms, societal value terms, dollars, . . .), the values of the outcomes can be compared and ordered.

On the assumption that the benefit of positive pressure is restricted to patients with hyaline membrane disease, net benefit can be expressed as the difference between the utility of administering treatment to affected babies and the utility of withholding treatment in the same class of infants:

(1) Net benefit $= U_{treat\ dis} - U_{no\text{-}treat\ dis}$

The cost may be expressed as the difference between the utility of avoiding treatment in

unaffected babies and the utility of administering treatment in the unaffected class of infants:

(2) Net cost $= U_{\text{no-treat no-dis}} - U_{\text{treat no-dis}}$

The expected value (EV) of the two courses of action may be calculated as follows:

(3) $EV_{\text{treat}} = (p)\, U_{\text{treat dis}} + (1-p)\, U_{\text{treat no-dis}}$

(4) and $EV_{\text{no-treat}} = (p)\, U_{\text{no-treat dis}} + (1-p) U_{\text{no-treat no-dis}}$

A rational choice is the course of action with the higher expected value. When the expected values of the two courses of action are equal, it is rational to remain indifferent. From the above expressions, the probability value at the indifference point is derived as follows:

Equation (3) = Equation (4)

and solving for p (the probability at the indifference point),

$$p = \frac{U_{\text{no-treat no-dis}} - U_{\text{treat no-dis}}}{U_{\text{treat dis}} - U_{\text{no-treat dis}} + U_{\text{no-treat no-dis}} - U_{\text{treat no-dis}}}$$

If the probability of the disease in a given patient is greater than the probability at the indifference point, treatment is rational; if the probability of the disease is less than the 'indifferent' value, withholding treatment would seem to be rational.
(Adapted from the arguments of S.F. Pauker and J.P. Kassirer of Tufts University)

*The utilities of positive-pressure treatment and no treatment must take into account the probability of survival with complications (e.g. blindness, brain damage, ...) under these two courses of action.

(p 135) for conditional probability (using prior information about the frequency of diseases and the accumulation of indicants).

The cost-benefit trade-off of therapeutic decisions may also be calculated by assigning numerical values to probabilities and to utilities.

Function of randomized trials When viewed from this structured perspective, the function of randomized clinical trials in medicine emerges in a clear light. The formal exercises provide the basis for predictions of the outcomes following specific treatments. The hard-to-obtain information is used by physicians in everyday management decisions for individual patients, but it should be understood that the translation of these medical judgments into practical actions requires that patients (or their surrogates) make value judgments.

Subjective judgments

The question of whether utilities should be regarded as belonging to the theory of probability may be argued at some length. Howard Raiffa of

> The player on the other side is hidden from us. We know that his play is always fair, just, and patient. But also we know, to our cost, that he never overlooks a mistake, or makes the smallest allowance for ignorance. To the man who plays well, the highest stakes are paid, with that sort of overflowing generosity with which the strong show delight in strength. And one who plays ill is checkmated—without haste, but without remorse.
>
> T.H. Huxley

Harvard University has pointed out that the subjectivists (those who hold with Reverend Bayes) wish to introduce intuitive judgments and feelings directly into the formal analysis of a decision problem. The objectivists argue that these subjective aspects are best left out of the formal analysis and should be used only to bridge the gap between the real world and the objective results obtained in the use of a formal model.

Despite the philosophical debates, there can be little doubt that a patient's subjective impressions of the state of the world cannot be ignored. And the analytic approach of decision theory forces us to make a distinction between the roles of the two participants in medical decisions: the doctor and the patient (or a surrogate).

We assume from observing ordinary human behavior in a world full of risks that rational persons act as if they had sets of probabilities and utilities that determine their choices of action at every turn in their lives. The principle of rationality is followed when there is a consistency of choices that increase the expectation of gain in utility. In a medical setting, it is reasonable to expect that when doctors provide patients with estimates of the probable outcomes of interventions, the sufferers will make rational decisions about the options that are available to them.

The seeds of misunderstanding lie buried in the phrase 'gain in utility'. Clearly, this function can only be defined and weighted in meaningful terms by patients and their families. Not surprisingly, private and public values are not always identical, and the value scales of the medical profession may be quite different from those of both patients and communities. The inevitable conflicts lead to some of the ethical problems that I will discuss in the next chapter.

Halting acceptance of formal methods

It would be misleading to leave the impression that physicians have adopted the reasoning of decision theorists to guide decision making at the bedside. The development of formal methods for the quantification of judgment to replace traditional intuitive approaches in medicine has been halting. The principal difficulty is that of making numerical estimates of the utilities of possible outcomes of medical interventions. I believe the resistance stems

from our reluctance to make what seems to be cold and scheming statements about the *prospect* of negative utilities associated with horrendous possibilities like blindness, brain damage, and death. On the other hand, the emotional blocks are surmounted when we are asked to make the same cost estimates after the fact: the numbers are supplied for tort claims and insurance awards.

Additionally, the sensitive question concerning the extent to which medical practice should be based on economic principles is not easily discussed. I believe, however, that we cannot ignore the experience in other fields concerned with our lives and well-being; there has been a fairly general shift from haphazard methods to disciplined systems for decision making. Moreover, the benefits of systematic approaches in medicine are testable. Improvement may be measured in formal comparisons by the gain in utility of the state of health of patients and by the cost of achieving a change.

WIDESPREAD MEDICAL ACTION

I have said that we may look to the results of randomized trials for guidance concerning probable outcomes of treatment in a specific patient and that we may estimate the cost of our errors in relying on this information. The potential cost mounts with the treatment of large numbers of additional patients; the price of wide extrapolation of a misleading result may turn out to be quite high (as we saw in the example of the national RLF trial). Thus, if the results of a bedside trial seem to be important, independent confirmatory trials are highly desirable.

Recommendations for repeated demonstrations

R.A. Fisher commented, 'In order to assert that a natural phenomenon is experimentally demonstrated we need, not an isolated record, but a reliable method of procedure ... we may say that a phenomenon is experimentally demonstrable when we know how to conduct an experiment which will rarely fail to give us a statistically significant result [i.e. a result in the same direction and of similar magnitude].' This caution was expressed many years ago in reference to agricultural studies in which the experimenter was able to exert much more control over extraneous influences than in the highly unstable conditions found in medical settings.

The need for replications of clinical trials was underscored by Mainland early in the history of the use of these formats: the only way to learn something about the safety of our numerical findings, he noted, is by more extensive exploration under other conditions, in other places, and at other times.

Evaluation of social effectiveness

An additional argument can be made for cautious, progressively staged extrapolation of the results of treatment trials rather than single-step widespread applications. The reasoning was advanced by Halfdan Mahler, the Director-General of the World Health Organization. He pointed out that health development is essentially a political and social process that should start off with the acceptance of the social function of health. Technical developments need to be applied in harmony with this social function.

In many communities the value of expensive treatments can be seriously questioned if measured in terms of their impact on improving the health status of the populations. Governments and people have a right to an objective assessment of medical innovations. This implies that critical evaluation of an intervention entails not only tests of effectiveness in modifying the course of disease but also controlled community trials to determine acceptance and to explore methods of implementing the new approach. The last step is one of evaluating social effectiveness.

Activism versus the slow pace

I have now proposed many roadblocks to the speedy and widespread use of the potential cures developed by modern biomedical research. But it must be obvious that these overly cautious views are not shared by all who are engaged in the search for ways and means to reduce human suffering. In fact, there is a fairly sharp division between two ideologies. Sackett has dubbed the adversaries as the 'evangelists' and the 'snails'.

Evangelists In the presence of an on-going toll of disability and untimely death, the evangelists conclude that pre-existing evidence plus common sense demand intervention *now* even in the absence of experiments to test whether extrapolations from optimistic pre-clinical studies will, in fact, alter risk.

This attitude is bolstered by public clamor for rapid dissemination of new biomedical knowledge, and for conversion of this information into forms directly useful in the control of major diseases. Legislators demand a visible payoff on the investment of public funds in medical research. The activist viewpoint has been called the 'Lyndon Johnson doctrine', after a speech in which he chided the American research community for its alleged failure to translate the findings of laboratory investigation speedily into practice.

Additionally, the evangelistic view tends to see rational decisions as 'terminal': a particular hypothesis is to be acted upon without further experiments.

Snails The methodologists are the snails. They are wed to a series of criteria and strategies that, if rigorously applied, will increase the likelihood

that the ultimate action is 'correct'. They insist that interventions must meet scientific and political criteria before they are widely implemented. And this school of thought is inclined to see rational decisions as 'sequential': a sequential decision is a judgment to perform further particular experiments.

Modern dangers Paradoxically, one can make a stronger argument for the snail's pace in the present era of rapid increase in the volume of biomedical information than in the leisurely past. This incongruity arises from the fact that the magnitude of the consequences of error has escalated so sharply in recent years. The harmless nostrums of the past have been replaced by exceedingly powerful agents, and the means for broadcasting information about treatments have become very efficient.

The image of the physician as a romantic warrior armed with a sword to conquer disease is outdated. The modern doctor may be likened to a weapons specialist manning the push-button console of a missile launcher. Modern conditions call for cautious tactics: limited forays and carefully planned battles against ignorance concerning medical matters.

Simple concern for the public interest dictates that we cannot ignore Claude Bernard's maxim: science teaches us to doubt and, in ignorance, to refrain.

Abatement of loss

The wary approach to choice under conditions of uncertainty is based on a consideration of the worst that can happen. One principle that has been suggested is called the 'minimax average loss rule': pick that strategy for which the largest average loss is as small as possible.

Patients must consider not only how often they may 'lose', but the magnitude of the loss. In gambling terms they must weigh the size of the risk against the size of their 'bankroll'.

The thalidomide disaster The issues were highlighted by the thalidomide tragedy that began when this drug was introduced in West Germany in 1957. It was regarded as a safe and useful medication, especially in the treatment of nausea during pregnancy. In late 1961, the West German Minister of Health issued a statement warning pregnant women not to take the drug; it was found to be associated with malformations (seal-like deformities of the limbs) that occurred in thousands of babies born to treated women throughout the world. Although the agent was not released for general use in the United States, 2½ million tablets were made available to 1267 American physicians for informal clinical 'investigation'.

Legal brake on drug certification As a direct result of the thalidomide experience, the Kefauver-Harris Amendment to the US Pure Food and Drug Laws was passed in 1962. For the first time legal regulations set out the formal steps to be taken in testing the safety and efficacy of a new drug for human use.

This law played an important role in damping the exuberance that characterized the proliferation of new drugs in the years following the spectacular successes of penicillin and polio vaccine. The rate at which new drugs were released for general use slowed abruptly, and there has been much debate about these developments. Many have concluded (especially as the memory of the thalidomide tragedy fades) that the restrictive regulations must be relaxed because they hamper the development of 'effective' drugs and slow medical progress. An advisory committee examined the charges against strict regulations in the United States and found that there had been a worldwide slowing of new drug development. The advisors concluded, however, that deceleration was 'mainly the result of limitations of scientific understanding of biological actions in disease processes'.

Phases of drug evaluation in the United States

Food and Drug Administration Guidelines

Phase I *Initial administration of a drug to human subjects.*
Following the satisfactory completion of studies in animal models, a new drug is administered to a few volunteers (usually healthy persons) in order to determine the metabolic transformation and excretion of the compound, and to estimate levels of tolerance. This step is followed by dose-ranging studies for safety and, in some cases, pilot trials of efficacy are carried out in selected patients.

Phase II *Early controlled clinical trials designed to evaluate efficacy and relative safety.*
Relatively small numbers of patients are enrolled in closely monitored studies (the numbers seldom involve more than 100–200 individuals who receive the new agent). Drugs considered to be effective and safe then enter Phase III.

Phase III *Expanded controlled and uncontrolled trials.*
These studies are performed after efficacy has been tentatively established. The stage is focused on gathering additional evidence of efficacy, safety, and tolerance; and on the definition of adverse effects. Hundreds, sometimes thousands, of patients are enrolled in Phase III studies conducted in outpatient clinics and hospital wards (an effort is made to approximate the environments in which the new agent will be used). Drugs which complete Phase III studies satisfactorily are released for use by practising physicians.

Phase IV *Post-marketing clinical trials.*
Additional studies to investigate the frequency rates of adverse reactions and specific pharmacologic effects are conducted in the field. Large-scale, long-term surveillance to determine the effect of a drug on morbidity and mortality is carried out in some instances to supplement and confirm pre-release data.

Coordinated advance in medicine

From the arguments that I have made in this book it should be clear that I cast my lot with the snails. If more attention is paid to the clarification of goals than to the perfection of means, the rate of innovative change in medicine will, indeed, be slower than it has been in the recent past. In the words of Lewis Mumford, 'A pace of change might [be] established in relation to human need. Instead of rapid advances, on the basis of uncoordinated knowledge in specialized areas, there [will be] the possibility of a slower, but better coordinated advance that [does] justice to the processes, functions, and purposes of life.'