

Universität Bielefeld/IMW

**Working Papers
Institute of Mathematical Economics**

**Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung**

Nr. 82

Hans W. Gottinger

The Simple Economics of Screening Programs:
An Application of Decision Analysis to Medical Screening.

May 1979



H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung
an der

Universität Bielefeld

Adresse / Address:

Universitätsstraße

4800 Bielefeld 1

Bundesrepublik Deutschland

Federal Republic of Germany

The Simple Economics of Screening Programs: An Application of
Decision Analysis to Medical Screening

Hans-W. Gottinger
University of Bielefeld

In the extensive literature on screening and prevention programs it is always implicitly assumed that selective screening on high risk patient groups, for those with coronary disease or breast cancer, leads to a significantly higher detection of true positives, and, therefore, entails a corresponding increase of the number of expected lives saved. Consequently, it is argued that this justifies increased costs of screening programs and related medical care. In this paper the value of a screening program is derived on the basis of decision analysis using as a single criterion mortality costs. The conclusion drawn suggests that increased screening costs could only be justified up to a certain qualified limit, but are not justified beyond this limit.

Suppose you consider two disease complexes A and B, A is a very serious disease, requiring careful monitoring and possibly elective surgery. B is far less serious that needs no further treatment but reveals similar symptoms as A. A screening program is defined as a set of tests conducted on the patient or class of patients that serves in finding unique identifications for patients having disease A.

The physician is considered to be a decision-maker or problem-solver who, to the best of his knowledge and to the availability of given medical technology, structures the decision situation in such a way that the best option is the one which minimizes expected mortality or morbidity considered as expected costs in the overall problem. Additional criteria could be meaningfully taken into account, and they would involve a weighting of multi-criteria, but, for the sake of simplicity, we concentrate only on the unique criterion of mortality.

The screening program, consisting of a set of diagnostic tests, often applied sequentially, is considered to be a detection device for finding out whether the patient has disease A or B. We use 'tests' here in a general sense. They might consist of patient history, physical findings or laboratory results. They may be presented in a form suitable for computational purposes, see Ledley and Lusted (1961). In general, this detection device is imperfect, i.e. error-bound, so that we are left with asking questions about the reliability of the test(s). For this purpose we could set up a Test-Reliability

Matrix for our simple problem

Table 1.

Test results \		Disease in Patient	
		A	B
A	A	True positive	False positive
	B	False negative	True negative

Here the entries in this table, read along the first or second row, respectively, could be interpreted as :

First row : 'Test results indicate disease A and the patient has disease A'

'Test results indicate disease A and the patient has disease B'

Second row: 'Test results indicate disease B and the patient has disease A'

'Test results indicate disease B and the patient has disease B'

Since by nature of the problem A is far more serious than B, ensuing a substantial higher cost in terms of mortality or morbidity therefore requiring immediate action on the physician's side, the Test-Reliability Matrix reflects this view in the naming of the entries. (In the case A has been ruled out, one could set up another table comparing B with C, etc. so that the disease with highest priority, requiring most medical attention, is taken proper care of.) Of particular concern here are those patients who on the basis of the test results will be treated on the false disease (false positives) and those for whom the results missed the true disease (false negatives).

Suppose then on the basis of these tests, for diseases A and B, completely different therapies, T_1 and T_2 , are suggested, for instance T_1 may involve surgery for constraining stomach cancer, T_2 is a mild drug treatment for treating a nonmalignant tumor. Suppose further we have sufficient statistical evidence on therapies T_1 and T_2 with regard to mortality costs, we could set up an outcome table on the decisions (costs) of treatment T_1 and T_2 .

Table 2. Outcomes of decisions (costs) measured by mortality per 1000 patients

Treatment \		Disease	
		A	B
T_1	A	1.50	.70
	B	30	0

The numbers in the entries, used here only for illustrative purposes, are collected statistically for a sufficiently long period of time. But for more practical reasons it might be advisable to decompose the data according to patient groupings pertaining to age or specific environmental conditions. Different groups such as old vs. young, or men vs. women may have quite distinct mortality costs, and the overall aggregate average cost matrix may not be applicable to group specific circumstances. For collecting enough group specific, disaggregate data we may run into difficulties of sufficient data acquisition. In that case we may apply advanced statistical techniques (i.e. multiple regression analysis) for overcoming these difficulties.

For any action, T_1 or T_2 , the average mortality cost, given as the expected value of mortality, can be computed after specifying the probability of each disease state, A or B. Unless one has a sufficient data base one often finds it difficult to calculate the probability of the disease state. In this case the physician is required to make an introspective judgment or reasonable guess and to come out with a subjective probability reflecting his professional judgment. Various methods to attain a subjective probability can be applied, see De Finetti (1972) or Gottinger (1979), in terms of betting quotients, or comparing disease states with events for which well-known (objective) probabilities exist.

Suppose the physician's prior probability of disease state A is $P(A) = .05$, and for B it is $P(B) = .95$. Then, on the basis of Table 2, we compute the expected value of decision:

$$EV(T_1) = .05 \times 1.5 + .95 \times .70 = .075 + .665 = .74$$

$$EV(T_2) = .05 \times 30 + .95 \times 0 = 1.5$$

Clearly, the action with lowest average mortality cost is best.

The computations show that if you want to make a terminal decision or equivalently, if the costs of gaining information about specifying P exceeds the benefit of this information, for patients with .05 probability of having disease A it is better to apply T_1 than T_2 .

In fact, looking at the average mortality costs, T_1 costs .74/1000 in mortality, but T_2 costs 1.5/1000 in mortality. This is so because the consequence of applying T_2 if A is true is severe enough to outweigh its small probability. (The underlying assumption is that 'waiting three months' or 'adopting a mild drug treatment plan' substantially decreases those patients' survival chances with disease state A).

As we can summarize, at this point, the best decision depends on the probabilities of the various disease states and on the costs of mortality associated to the given diseases. By fixing the mortality rates, one can easily determine the threshold probability at which point it becomes advisable to switch from strategy T_2 to strategy T_1 . The threshold probability can be calculated as follows:

$$EV(T_1) = P \times 1.5 + (1-P) \times .70 = .70 + .80P$$

$$EV(T_2) = P \times 30 + (1-P) \times 0 = 30P$$

$$\begin{aligned} \text{Equalizing, } EV(T_1) &= .70 + .80P &= EV(T_2) &= 30P, \\ &\text{yields } P &= .024 \end{aligned}$$

Costs of mortality depending on threshold probabilities

Cost: mortality /1000

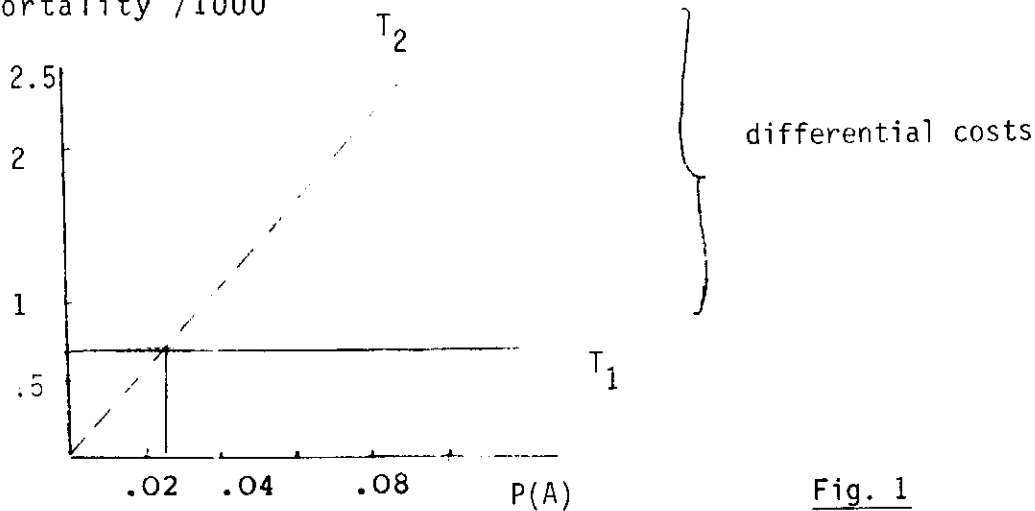


Fig. 1

One can simply interpret this figure along the following lines. We graph the mortality costs of T_2 (dashed line) and T_1 (solid line) as a function of the probability of A. If there is no chance of A occurring, there is no cost to T_2 , but there is always a cost to the more severe treatment T_1 . (If T is not considered as a treatment, but as yet another test it could be understood that applying the test itself affects the mortality rate. Suppose that in the case of breast cancer, screening for women mammography is generally applied. Then the additional risk for applying mammography for a particular patient group is reflected in its impact by increased mortality costs and therefore could be fit into this framework).

But to the extent that A becomes more likely, the ' T_2 -strategy' quickly becomes more dangerous. Up to the threshold level one should adopt a 'wait strategy', above this level one should switch to a more radical

therapy. Costs and probabilities obviously are interdependent. If you vary the costs in terms of probability, you correspondingly shift the threshold probabilities.

Recalling that a screening program consists of a sequence of tests to be performed, it appears as common sense reasoning, that under ideal circumstances a diagnostic test should show those persons or population groups to have the disease who actually have the disease. In other words, the best that any screening program can do is to correctly classify all patients. Since this requirement can be fulfilled only under exceptional, ideal circumstances, we could consider this state as our reference system and set out to enquire about the costs that obtain in such a system. It is clear that the costs in such a system cannot be decreased, given the present level of medical technology and medical knowledge. Since the value of a perfect screening program is the one that diagnoses a high risk factor group correctly and consequently leads to an adequate therapy, one would be interested in the value of information that minimizes diagnostic mistakes emanating from any screening program, coming close to a perfect screening program.

In other words, by improving the diagnostic-treatment situation one asks how much is more information worth?

By referring to Table 2 a perfect screening program yields mortality costs by computing

$$EV(T_1 | (\text{perfect screening})) = P \times 1.50 + (1-P) \times 0 = .05 \times 1.50 + .95 \times 0 = .075$$

as compared to

$$EV(T_1 | (\text{no screening})) = .05 \times 1.50 + .95 \times .70 = .74$$

In verbal interpretation, applying T_1 in case of perfect screening kills only .075/1000 of the population, an unavoidable cost, whereas applying T_1 given no screening kills .74/1000. This means that the cost of action in the light of perfect information is roughly ten times less than the cost associated to the best action on undifferentiated patients (apply T_1 to everyone). The expected value of perfect information, therefore, is equal to the difference of these two, since the mortality cost of .075/1000 appears unavoidable unless better methods of treatment are available.

To emphasize the point of optional treatment with or without screening on patient groups, we shall rewrite costs for each disease state in terms of regrets caused by mistakes. (To set up a regret matrix is a familiar procedure in statistical decision theory).

Table 3. Outcome expressed as differential mortality cost per 1000 (regret) due to improper treatment

Treatment	Disease in patient	
	A	B
T ₁	0	.70
T ₂	28.50	0

The number 28.50 in the lower left entry of the matrix comes from 30/1000, the costs of a 'wait strategy', minus 1.50/1000, the costs of the only correct treatment, T₁. The upper left hand entry gets zero, since the action is correct. The upper right entry remains as it is, since the cost of the correct action, T₂, (to be deducted) is zero.

Up to now we considered only the value of perfect screening as compared to no screening at all, taking the unavoidable mortality costs of a correct treatment as a basic reference point. The situation where perfect information can be acquired is rare, whereas partial information is often obtainable. Nevertheless, the expected value of perfect information is useful because it provides an upper bound to that for partial information. Therefore, more generally, we could exhaust the whole spectrum on evaluating different screening programs and figuring out the value of information in terms of mortality costs. It should be obvious that a crucial point in comparing these programs is the validity of their test results, that is the degree of accuracy according to which these tests identify and classify the correct disease-state patients.

Consider only a possible result of test validation for illustrative purposes.

Table 4.

Test results \ Test validation	Disease in patient	
	A	B
A	.8	.1
B	.2	.9

Such numbers might be obtained by collecting data on a series of patients who are all given the test, and then later investigated to see whether they really had the disease or not. Unless a representative sample of them can be autopsied, there may be problems with this test verification.

(The above table suggests that sensitivity of the test is .8, and specificity of the test is .9).

Now, how do test results affect estimates of the probability of disease? Bayesian statistics provide techniques for revising initial or prior probabilities in the light of new information. The information must be new to have any effect. In statistical analysis, the events are said to be independent if information about the occurrence of one event does not change our estimate of the probability that the other event occurred.

Independence of medical evidence is hard to judge, to estimate whether two tests are really independent in their predictions, one may have to collect substantially more cases. Sometimes there may be theoretical reasons to believe tests are independent - one may be biochemical, and another anatomical. Bayes' theorem weights prior probabilities by their likelihood, it follows from the definition of conditional probability. The conditional probability of A given that the test indicates A is defined by

$$P(A|Test = A) = \frac{P(A \text{ and } Test = A)}{P(Test = A)},$$

i.e. is defined to be the probability of both A and a true A-test result divided by the total probability of a positive A test result. In the numerator, we see the effects of our hypothetical test on patients .05 of whom are assumed (a priori) to have A. The test identifies 80% of them correctly, so .05 x .8 = .04 are identified correctly as having A by the test. Also in the denominator are the 10% of B

falsely called A by the test. The test will say that $.05 \times .8 + .95 \times .1 = .135$ have A, and $.04/.135 = .296$ of those so identified really will have it.

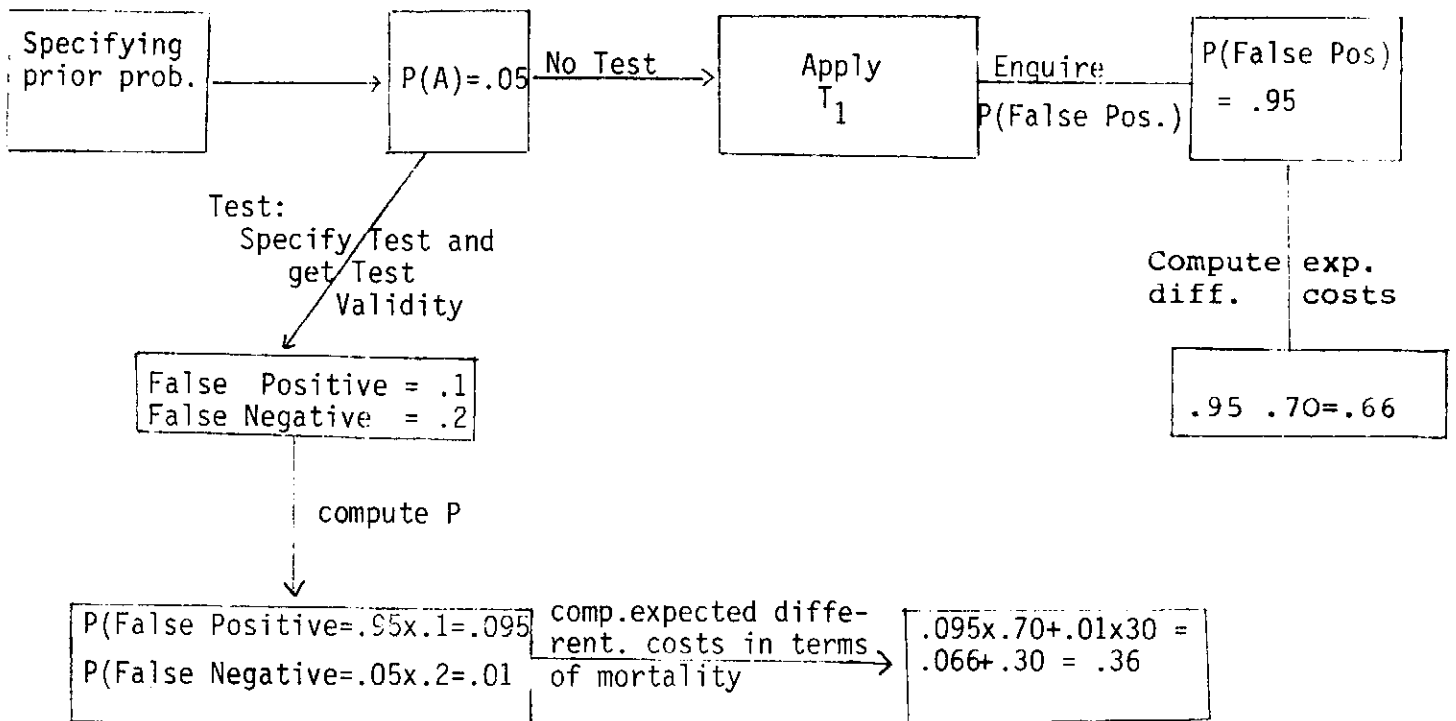
A similar calculation shows, namely

$$P(A|Test = B) = \frac{P(A \text{ and Test} = B)}{P(\text{Test} = B)} = \frac{.05 \times .2}{.05 \times .2 + .95 \times .9} = .012$$

that only .012 of those the test says have B will have A. The test has split the formerly homogeneous group of patients, each of whom pretended to have a 5% chance of A, into two distinct groups. One group has an almost 30% chance of A, and the other about a 1 % chance of A.

Let us see how we could assemble the various bits of information contained in Bayes' theorem: the differential cost (regret) matrix and the test validity data, forming the likelihood, to construct the value of a test represented only by a single criterion, the mortality rate. In general, the value of the test is simply computed by subtracting the average costs of the best action before the information of the test is available, from the best action afterwards.

Table 5. Flow-chart



These computations presented in a flow-chart form use the added costs of mistakes over correct actions in the differential cost matrix.. Applying T_1 immediately kills .70/1000 of those having B. Given the prior probability on B the expected additional mortality without the test amounts to $.095 \times .70 = .66$. After the test only 10% of these are subject to T_1 -treatment, whereas 20% of the true A cases wait, for a differential cost of .36. The test has saved exactly .30/1000 deaths. Unless the test itself kills more than that, it is justified by using reduction of mortality as a single criterion. (It may not be justified by a different criterion, e.g. resource costs associated to the test but this could be determined only by an appropriate benefit cost calculation of the test.)

In principle, the same computations go through in a sequence of tests making the entire screening program. Suppose one wants to improve the results in a further reduction of mortality by taking the latest test as a reference point. The only thing that changes is to replace the prior probabilities by posterior probabilities - computed according to Bayes' theorem on the basis of the previous test results. Of course, one has to take care of a strict statistical independence assumption by conducting the tests. If it turns out, say, that the subsequent test yields a differential cost of .20, as compared to the first test of .36, then the subsequent test is worth at most .16 in mortality. Suppose, then, without loss of generality, a screening program consists only of these two tests. Then the value of the screening program is the sum of the value of each of the two tests, applied independently and sequentially, e.g.
 $.30/1000 + .16/1000 = .46/1000$.

If besides mortality, there is a serious consideration by the medical decision maker of taking into account resource costs generated by the screening program, then this could be achieved by letting the consumer (patient) determine his willingness to pay for the unit monetary cost of the test in exchange for a reduction of mortality. From decision theoretic principles we know that if someone has a 1 in 1000 chance of having a life-threatening problem, he might be willing to pay a certain amount to find out.

In fact, J.P. Acton (1973), by using this methodology as originally proposed by T.C. Schelling (1968), worked out a refined catalogue of questionnaires that inquired about people's preferences with regard to various public programs including a screening-monitor-pretreatment program for heart diseases.

However, it appears as an immediate problem that consumers of medical care may not adopt this obviously rational approach, anxiety may prevent them to find out whether there is a medical problem that can be detected by screening.

In the sequel we deal with some important extensions and complications of the previous analysis.

(1) Test results out of screening programs may not be split into two categories, but instead may discretely range over several levels. Readings of biochemical levels might give rise to several, non-unique interpretations. The picture interpreted by a radiologist may present convincing, weak, confusing or no evidence of the disease in question. Suppose that the information in the patient's history and physical findings have been grouped into five categories, as shown in the next table.

Table 6.

Test results	Disease	
	A	B
A	.2	.1
A likely	.3	.1
Non-conclusive	.3	.3
B likely	.1	.2
B	0	.3

For instance, the category 'A likely' contains 30% of A cases and 10% of B cases.

If a test could provide that much of more detailed, refined information it should not be arbitrarily calibrated to two categories, calling the top two categories A and the bottom three B. It is important to report test information as precisely and completely as possible.

(2) As indicated previously, costs in health programs have a multi-criteria representation, the investigation of single components may be only of limited value. For some diseases health costs are fairly easy to evaluate, long-range aftereffects are not too important. For cancer and many other diseases, other health costs must be considered. Suppose a breast cancer screening program screens women at 50, and saves 10 people who otherwise would have died at 55, extending their lives to 70.

On the other hand, suppose the radiation of testing causes 10 additional cases of cancer so that 10 people who would otherwise have died at 70, die instead at 65. The net change in mortality of the program is zero, but the net gain in years of life is 100. In such case you feel that years-of-life is a better measure of health costs than simple mortality. For cancer, in particular, quality-of-life is important. Healthy years are rated higher than low-quality years.

We can minimize either average immediate mortality costs, or average years-of-life costs or average resource costs. The best cost-minimizing action is different for different types of costs. Not all costs can be minimized at the same time. Thus, even the best treatment plan involves tradeoffs of one type of cost for another. This seems to be in accord with pursuing 'compromising strategies' by implementing screening programs, and, in fact, this is proposed by some researchers in the field (see L.E. Blumenson, 1977).

(3) If various categories of costs such as resource costs, disability days (a surrogate measure of years-of-life costs) and mortality are plotted against threshold probabilities of disease state A, on a continuous scale between 0 and 1, the medical resource costs of a delayed T_1 -treatment could amount to being only twice as high as a timely T_1 -treatment, but according to Table 2 the mortality costs are 30 times as high. This is reflected by the fact that the probability of A for the minimum mortality costs is considerably lower than the probability point for the minimum resource costs. In other words, aggressive treatment costs money but saves lives.

In general, all the marginal cost curves rise sharply as the probability of A increases. A T_1 -strategy is definitely indicated for high probabilities. All the cost curves decrease starting from probability zero. Excessive T_1 -treatment at low probability of a serious disease is expensive and dangerous. (In fact, this seems to support empirical findings that excessive surgery at low probability of serious diseases is likely to cost lives besides eating up a substantial portion of resource costs).

Suppose the physician chooses to reduce the threshold probability slightly (e.g. applying a T_1 -strategy on patients with slightly lower probabilities of A), he could save (say) 100 more lives a year at a cost of 10,000,000 monetary units and 50,000 disability days. Each life thus costs 100,000 monetary units and 500 days. Is this an acceptable price for a life saved? Now, if you think life is worth more,

you should choose a threshold probability point closer to the minimum mortality point. If you think the costs are too high, you should use a somewhat higher probability point.

References

- J.P. Acton (1973), Evaluating Public Programs to Save Lives: The Case of Heart Attacks, Rand Corporation, R- 50-RC, Santa Monica, Ca.
- L.E. Blumenson (1977), "Compromise Screening Strategies for Chronic Disease", Mathematical Biosciences 34, 79-94
- B. de Finetti (1972), Probability, Induction and Statistics, Wiley: London, New York
- H.W. Gottinger (1979), Elements of Statistical Analysis, W. de Gruyter: Berlin, New York
- R.S. Ledley and L.B. Lusted (1961), "Medical Diagnosis and Modern Decision Making", Proc. of Symp. Appl. Math., Vol. XIV, Am. Math. Soc., Providence
- T.C. Schelling (1968), "The Life you Save may be your Own", in Problems in Public Expenditure Analysis (S.B. Chase, ed.), Brookings: Washington, D.C.

Cost-Benefit Methodology in Medical Care: A Critique
of Recent Studies

The application of cost-benefit methodology to medical care services has come a long way from its first cautious steps in the late 1950s and early 1960s (see, for example, Fein 1958; Weisbrod 1961; Klarman 1965). In the early studies, benefits were measured by the change in a discounted stream of earnings caused by a medical care intervention.

Beginning with Schelling's seminal paper in 1968, economists began to realize that earnings as a measure of benefits was seriously defective (see also Taylor 1969). Subsequently, there has been widespread agreement that the proper measure of benefit is given by the utility a well-informed consumer receives from the intervention; empirically, this can be represented by the amount the consumer would be willing to pay for the intervention (or the amount the consumer would accept to forego it; if the willingness-to-pay and willingness-to-forego differ, a specialist should be consulted). Unfortunately, there has been no large-scale attempt to implement a willingness-to-pay measure, and existing studies (for example, Acton 1973) remain under a cloud of suspicion that consumers did not understand the questions and so did not give valid responses (Raiffa, Schwartz, and Weinstein 1976). Decision-theoretic approaches to diagnosis and treatment are in the same spirit as the willingness-to-pay approach in that they attempt to ascertain the patient's preferences for certain outcomes (Schwartz, et al., 1975; Pauker 1976). Decision analysis is especially useful when uncertainty is an important part of the problem. This prologue is relevant to the papers on computed tomography (CT) (Larson, et al., 1977), and X-ray for lower back pain (Rockey, et al., 1977), and the cost of rheumatoid arthritis (Meenan, et al., 1977).

I am delighted to see the awareness of costs that these papers evidence; I can still vividly recall discussing some years ago a decision tree for laboratory testing that a physician had drawn up; the tree omitted any reference to cost. When I pointed out that cost really belonged in the tree, the physician became indignant, arguing that only risk and value of information were relevant in medical decisions. Times

have certainly changed!

It is important to be clear about why one is interested in cost. Ultimately, one wants to make the most of society's scarce resources. CT and X-rays use resources that could have been used to produce other goods and services (or to reduce "bads" such as pollution). Such resource use shows up in an accounting framework as cost. Hence, the question comes: Is this the best use that can be made of these resources, the best that can be done for this cost? Such a question cannot be answered unless we can value outcomes or measure benefits.

In the measurement of benefits, I am afraid that some authors are showing signs of a syndrome noted by Oscar Wilde to afflict economists, who were said to know the cost of everything and the value of nothing. In the case of the paper concerning CT, the authors find that measures such as length of hospital stay and speed of work-up do not differ before and after the introduction of CT, and they therefore judge CT solely on whether it raises or lowers diagnostic costs. For two of the three diagnoses that they examine (suspected brain tumor and suspected hydrocephalus), CT lowered cost; for the third (suspected cerebrovascular disease), it raised it. Are we to conclude that CT should not be used in patients with suspected cerebrovascular disease? The authors think so: "However, for patients with suspected cerebrovascular disease, by far the most common inpatient neurological diagnosis, our studies suggest that CT does not have a beneficial impact on patient care and that 'need' projections should not include patients with this common diagnosis."

Are the authors right? Perhaps so, but there are signs of Wilde's Syndrome here. The authors note that there was a significant decrease in lumbar punctures in these patients, from around 25 percent to 10 to 15 percent. Would an informed consumer be indifferent between receiving CT and a lumbar puncture? I doubt it, since one can get a prolonged headache from the lumbar puncture and occasionally paralysis or meningitis might even occur. The proper question to ask is: Are informed consumers willing to pay the additional cost of CT for the

convenience of not undergoing a lumbar puncture?¹⁾ If the answer is yes, then CT should be used to diagnose suspected cerebrovascular disease. Moreover, some patients may and some may not; Pauker finds that patients differ in their preferences for possible outcomes of coronary artery surgery (Pauker 1976). Thus, in many cases, it may not be possible to make general statements.

An issue of benefit valuation also occurs in the study of X-rays for lower back problems. In this case, the authors look at measures of mean days lost and symptom status in a group of patients that received a spine X-ray and a group that did not. While these measures showed no difference, about 15 percent of the group that did not receive an X-ray was either unsatisfied or saw another physician; less than half that figure was unsatisfied or saw another physician in the group that received an X-ray. The authors state: "From our data and a review of the literature, we conclude that back X-ray examinations have negligible diagnostic value in otherwise healthy patients under 50 years of age with nontraumatic backache ... We recognize that any clinical strategy which reduces the use of the back X-ray examination may require concomitant patient education to maintain patient satisfaction."

Again, the authors may be right, but again their data are not fully persuasive. First, the cost of the "concomitant patient education" needs to be considered, as well as the cost of any residual dissatisfaction or additional visits. But there is a more important issue. It may be the case that a spine X-ray could rule out or confirm certain quite infrequent problems for which intervention could make a difference if the problem were caught early. If the problem were sufficiently infrequent, it needn't show up in the sample of patients the authors examined. Physicians have suggested

1) A similar point can be made about the other two diagnoses studied; for them the frequency of pneumoencephalograms declined after CT, and I doubt that an informed consumer would be indifferent between CT and a pneumoencephalogram either. In these cases, however, CT also lowered cost.

that cancers, especially multiple myeloma, and bacterial osteomyelitis could be missed if an X-ray were taken, as could an abscess. They have also raised the issue of whether one can sufficiently trust the history to rule out trauma with certainty. The point here is simple: If I have a 1 in 1000 chance of having a life-threatening problem, I might be willing to pay \$ 32 to find out. (This would be in the range of values suggested by Weinstein and Stason [1976] for the value of a statistical life.) Note that there is also reassurance value in a (true) negative finding. From decision-theoretic principles, we know that X-ray is less likely to be preferred the lower is the probability that the X-ray could detect an important problem. Hence, the authors' conclusion that X-ray is of negligible diagnostic value in patients under 50 without trauma may be correct, and that group may be the optimal subset not to X-ray. But no analysis is presented to distinguish this particular group.

There is also an important point connected with the measurement of cost that both papers raise. One really wants to know what opportunities society forewent to use resources in a given way, the opportunity cost. There are two important implications for these problems: (1) Charges for ancillary services carry a notoriously high markup. As a result, use of charges will overstate the opportunity costs. (2) The appropriate cost figure is marginal cost; that is, what did the additional X-rays or additional CT cost? That is what society foregoes to deliver the service. But the charge may relate to average cost rather than marginal cost. For example, if the CT scanner were in place and being used for suspected brain tumor and suspected hydrocephalus, but were not fully utilized, the marginal cost of running an extra scan would be much less than the average cost. The charge would then further overstate the opportunity cost. For both these reasons, the studies appear to have overstated the relevant cost of the diagnostic services they consider.

In sum, it is important to be broad in one's definition of benefits; anything that is of potential value to the informed consumer is fair game. If possible, one should set up a me-

chanism to obtain measures of benefit from consumers themselves, and should allow for the possibility that different patients would prefer to be treated differently. If benefits include ascertaining the likelihood of low probability events, sufficiently large samples need to be used to make this possible. Measures of cost should reflect the marginal cost of the good or service, and should also reflect the true resource cost, rather than "charges." [Gottinger, 1979].

The third study differs markedly from the other two in that it focuses upon the costs imposed on certain individuals by rheumatoid arthritis. It notes quite correctly that costs include not only medical costs, but also reduced opportunities in the use of one's time. The authors should, but do not, distinguish between a loss in earning caused by a reduction in hours and a loss caused by a reduction in the wage rate. To the degree that the loss in earning comes as a reduction in hours, one gains time that one may use in other ways, and loss of earnings overstates the true loss. To the degree that the loss is in the wage rate, there is no compensation. One suspects that much of the earnings loss the authors show was in the wage rate, so that the authors' estimated costs may not be far off the mark.

Assuming this to be the case, the authors demonstrate that those with rheumatoid arthritis bear large costs, and raise the issue of whether nonmedical costs should be insured in addition to medical costs. The answer to this issue must consider a host of questions that are beyond the scope of this comment. These questions include:

- (1) To what degree are public and private disability insurance programs performing badly?
- (2) If public disability insurance is to be expanded, should it be financed by premiums (as group health insurance is), by payroll taxes, by general revenues, or by taxes related to the product generating the disability (for example, a tax on coal to finance a black-lung benefit)?

(3) How should disability benefits relate to previous earnings in such a program?

(4) What is the definition of a disability? Is it defined solely by a medical diagnosis such as rheumatoid arthritis? Must functional limitation be shown? If so, who decides what limitation is sufficient and by what criteria?

(5) How do we apply public disability insurance to housewives? Do we pick an arbitrary value for the housewife's services? Or, do we let the household buy different values of insurance, just as it now can buy differing amounts of life insurance for the housewife?

(6) How do we apply public disability insurance to members of the household other than those employed and housewives? If an elderly parent becomes disabled and comes to live with the child, this may mean one member of the household must stop working. Can that be an insurable risk? I do not propose to answer these questions, but only to suggest some of the complexities in this area.

From this vantage point, cost-benefit analysis and decision theory have many useful applications in medicine. Indeed, physicians should know rudimentary decision theory because of its relevance to decision-making in patient care, although they cannot be expected to acquire the detailed knowledge of these tools that an economist possesses. I detect increasing collaboration between economists and physicians in applying these tools to medicine. Such collaborations might prove useful for future progress.

REFERENCES

- Acton, Jan Paul, "Evaluating Public Programs to Save Lives: The Case of Heart Attacks," Santa Monica: The Rand Corporation, R-950-RC, January 1973.
- Fein, Rashi, Economics of Mental Illness, New York: Basic Books, 1958.
- Gottinger, H.W., "The Simple Economics of Screening Programs: An Application of Decision Analysis to Medical Screening," Working Paper 'Institute of Mathematical Economics', University of Bielefeld, 1979.
- Klarman, Herbert E., "Syphilis Control Programs," in Measuring Benefits of Government Investments, ed. Robert Dorfman, Washington, D.C.: The Brookings Institution, 1965.
- Larson, Eric B., et al., "Impact of Computed Tomography on Patient Care: Relevant Information for Efficient Use of Health Care Resources," proc. of the Robert Wood Johnson Clinical Scholars Program Annual Meeting, Lincolnshire, Ill., October 27, 1977.
- Meenan, Robert F., et al., "The Costs of Rheumatoid Arthritis: A Patient-Oriented Study of Chronic Disease Costs," proc. of the Robert Wood Johnson Clinical Scholars Program Annual Meeting, Lincolnshire, Ill., October 27, 1977.
- Pauker, Stephen, "Coronary Artery Surgery: The Use of Decision Analysis," Annals of Internal Medicine, 85, July 1976, pp. 8-18.
- Raiffa, Howard, William B. Schwartz, and Milton C. Weinstein, "Evaluating Health Effects of Societal Decisions and Programs," in Decision-making in the Environmental Protection Agency: Selected Working Papers, Vol. 11b, National Academy of Sciences, Washington, D.C., 1977, pp. 1-81.
- Rockey, Paul H., et al., "The Usefulness of X-ray Examinations in the Evaluation of Patients with Back Pain," proc. of Lincolnshire, Ill., October 27, 1977.
- Schelling, T.C., "The Life You Save May Be Your Own," in Problems in Public Expenditure Analysis, ed. Samuel B. Chase, Jr., Washington, D.C.: The Brookings Institution, 1968

- Schwartz, William B., et al., "Decision Analysis and Clinical Judgment," The American Journal of Medicine, 55, October 1973, pp. 459-484.
- Taylor, Vincent D., "How Much is Good Health Worth?" Santa Monica: The Rand Corporation, P-3945, July 1969.
- Weinstein, Milton C., and William B. Stason, Hypertension: A Policy Perspective, Cambridge: Harvard University Press, 1976.
- Weisbrod, Burton A., Economics of Public Health: Measuring the Economic Impact of Diseases, Philadelphia: University of Pennsylvania Press, 1961.

WORKING PAPERS

- Nr. 1 C.C. v. Weizsäcker: The Political Economy of Stability in Western Countries - The Wicksell Lectures - Stockholm, 2nd and 4th May, 1972, publ. at Almqvist and Wiksell, Stockholm 1972
- Nr. 2 C.C. v. Weizsäcker: Modern Capital Theory and the Concept of Exploitation, July 1972, publ. in Kyklos, May 1973
- Nr. 3 C.C. v. Weizsäcker: A New Technical Progress Function (1962), August 1972
- Nr. 4 C.C. v. Weizsäcker: Long Term Global Optimization in Educational Planning. A Simple Example, September 1972
- Nr. 5 C.C. v. Weizsäcker: Intergenerationelle Einkommensverteilung: Einfache Beispiele für Wirkungen steuerlicher Maßnahmen und für optimale Steuerstruktur, October 1972, publ. in Neue Aspekte der Verteilungstheorie, G. Bombach, B. Frey, B. Gahlen(eds.), Tübingen 1974
- Nr. 6 C.C. v. Weizsäcker: Kenneth Arrow's Contribution to Economic Sciences, November 1972, publ. in Swedish Journal of Economics, Vol. 74, 1972
- Nr. 7 C.C. v. Weizsäcker: Morishima on Marx, January 1973, revised April 1973, publ. in Economic Journal, December 1973
- Nr. 8 Reinhard Selten: A Simple Model of Imperfect Competition, where 4 are Few and 6 are Many, February 1973, publ. in International Journal of Game Theory, December 1973
- Nr. 9 C.C. v. Weizsäcker: Notizen zur Marxschen Wertlehre, April 1973
- Nr.10 Hans W. Gottinger: Some Measures of Information Arising in Statistical Games, May 1973
- Nr.11 Sergiu Hart and Elon Kohlberg: Equality Distributed Correspondencies, July 1973
- Nr.12 C.C. v. Weizsäcker: Grenzen der traditionellen Globalsteuerung, October 1973, publ. in Jahrbücher für Nationalökonomie und Statistik, Band 189, Heft 1/2, 1975
- Nr.13 C.C. v. Weizsäcker: Reply to Mrs. Robinson, Morishima and Wolfstetter, November 1973
- Nr.14 C.C. v. Weizsäcker: Substitution Along the Time Axis, publ. in Kyklos, Vol. XXVII, 1974
- Nr.15 John C. Harsanyi: The Tracing Procedure: A Bayesian Approach to Defining a Solution for n-Person Noncooperative Games, Part I, May 1974
- Nr.16 John C. Harsanyi: The Tracing Procedure: A Bayesian Approach to Defining a Solution for n-Person Noncooperative Games, Part II, May 1974
- Nr.17 John-ren Chen: Produktion, Konsum und Markt des Nahrungsmittels Reis in Taiwan: Ökonometrische Untersuchung und Prognose, May 1974
- Nr.18 Reinhard Selten: The Chain Store Paradox, July 1974
- Nr.19 C.C. v. Weizsäcker: Political Limits of Traditional Stabilization Policy, June 1974
- Nr.20 Reinhard Selten: Bargaining under Incomplete Information - A Numerical Example, July 1974
- Nr.21 John C. Harsanyi: Nonlinear Social Welfare Functions or Do Welfare Economics have a Special Exemption from Bayesian Rationality? July 1974
- Nr.22 John C. Harsanyi: A Further Note on Rawls's Theory, July 1974
- Nr.23 Reinhard Selten: Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, August 1974
- Nr.24 John-ren Chen: Abwertungs- und Aufwertungseffekte in einer Volkswirtschaft mit besonderen Außenhandelsverflechtungen, Sept. 1974
- Nr.25 Hans W. Gottinger: Existence of a Utility on a Topological Semigroup, January 1975

- Jr. 26a Hans W. Gottinger: Decomposition for a Static Stochastic Hierarchical System, February 1975
- Nr. 26 International Workshop on Basic Problems of Game Theory, Bad Salzflun, September 2 to 17, 1974: COLLECTION OF ABSTRACTS
- Nr. 27 Hans W. Gottinger: Decomposition for Stochastic Dynamic Systems Part I, April 1975
- Nr. 28 Claude Hillinger: Real Income, The Cost-of-Living and Consumer Surplus: A Unified Approach to their Conceptual Foundation, April 1975
- Nr. 29 Lawrence H. Nitz: Machiavellianism and Competitive Social Contacts in a Limited Information Convention Game, May 1975
- Nr. 30 Hans W. Gottinger: Decomposition for Stochastic Dynamic Systems Part II, May 1975
- Nr. 31 Hans W. Gottinger: Sequential Analysis and Optimal Stopping, May 1975
- Nr. 32 John-ren Chen: The World Cotton Market (1953-1965): An Econometric Model with Applications to Economic Policy, May 1975
- Nr. 33 Hans W. Gottinger: Some Applications of a Result in Control Theory to Economic Planning Models, June 1975
- Nr. 34 Lawrence H. Nitz: Ameliorative Coalition Behavior and Individual Strategies, August 1975
- Nr. 35 John-ren Chen: The Effects of an Increase of the Energy Price on Macroeconomic Activity: A Comparative Static Approach, August 1975
- Nr. 36 Klaus Binder: Portfolio Selection: Empirische Untersuchung für die Bundesrepublik Deutschland, August 1975
- Nr. 37 Hans W. Gottinger: On a Problem of Optimal Search, October 1975
- Nr. 38 Hans W. Gottinger: Lecture Notes on Concepts and Measures of Information, October 1975
- Nr. 39 A. Hoqqatt, R. Selten and D. Crockett, S. Gill, J. Moore: Bargaining Experiments with Incomplete Information, November 1975
- Nr. 40 T. Marschak and R. Selten: Convolutions, Inertia Supergames, and Oligopolistic Equilibria, November 1975
- Nr. 41 Dierk Bitter: The Kernel for the Grand Coalition of the Four-Person Game, January 1976
- Nr. 42 John-ren Chen: Ein Makroökonomisches Modell für Taiwan, I-V, May 1976
- Nr. 43 John-ren Chen: Ein Makroökonomisches Modell für Taiwan, VI - X, May 1976
- Nr. 44 Peter Rice: A Solution Theory for the Finite Negotiation Problem May 1976
- Nr. 45 Reinhard Selten: A Simple Game Model of Kidnapping, June 1976
- Nr. 46 Reinhard Selten: The Equity Principle in Economic Behavior, July 1976
- Nr. 47 Siegfried Brune: On the Regions of Linearity for the Nucleolus and their Computation, August 1976
- Nr. 48 Siegfried Brune: Computation of the Nucleolus for Superadditive 4-Person Games, September 1976
- Nr. 49 Peter Rice: A Note on the Hicks Theory of Strike Bargaining, September 1976
- Nr. 50 Alfred Zauberman: Note on Assimilation of Optimal Control Theory and Ramifications in Economic Planning Theory, September 1976
- Nr. 51 Alfred Zauberman: Planning under Indeterminacy, December 1976
- Nr. 52 John-ren Chen: Die Entwicklung der "Terms of Trade" der Entwicklungsländer von 1960 bis 1975, January 1977
- Nr. 53 Hans W. Gottinger: Simple Decision Procedures - An Expository Paper, February 1977
- Nr. 54 R.W. Scholz: Sequential Two Person-Zero-Sum-Games with Incomplete Information and Incidental Sequence of Moves, June 1977

- Nr. 55 Reinhard Selten: A Model of Oligopolistic Size Structure and Profitability, September 1978 (revised)
- Nr. 56 John-ren Chen: Ein Außenhandelsmodell für die "Hard-Core"-Entwicklungsländer von 1960 - 1975, July 1977
- Nr. 57 Rolf Stoecker: Altruism and Performance in Bertrand-Duopoly-Experiments, July 1977
- Nr. 58 Reinhard Selten und Werner Güth: Macht Einigkeit stark? - Spieltheoretische Analysen einer Verhandlungssituation, August 1977
- Nr. 59 Hans W. Gottinger: Complexity and Social Decision Rules, September 1977
- Nr. 60 K.-P. Kistner und R. Subramanian: Regenerative Eigenschaften von Modellen der Zuverlässigkeitstheorie, October 1977
- Nr. 61 Reinhard Selten: Coalition Probabilities in a Non-Cooperative Model of Three-Person Quota Game Bargaining, November 1977
- Nr. 62 Hans W. Gottinger: A Markovian Decision Process with Hidden States and Hidden Costs, November 1977
- Nr. 63 Reinhard Selten und Werner Güth: Risikodominanz in einem Markteintrittsspiel, November 1977
- Nr. 64 John-ren Chen: The Effects of Stabilization Policy under Controlled Flexible Exchange Rates, January 1978
- Nr. 65 John-ren Chen: The International Transmission of Business Cycles, January 1978
- Nr. 66 Rolf Stoecker: Schlußeffekt in einem Preisduopolexperiment, February 1978
- Nr. 67 John-ren Chen: Prüfung der technischen, ökonomischen und politischen Voraussetzungen für den Abschluß eines internationalen Rohstoffabkommens mit Bufferstock, Mindest- und Höchstpreisen zur Stabilisierung des Weltmarktes für Baumwolle und Baumwollgarne, February 1978
- Nr. 68 Hans W. Gottinger: Markovian Decision Processes with Limited State Observability and Unobservable Costs, April 1978
- Nr. 69 Reinhard Selten und Werner Güth: Game Theoretical Analysis of Wage Bargaining in a Simple Business Cycle Model, September 1978
- Nr. 70 Reinhard Selten: Limited Rationality and Structural Uncertainty, September 1978
- Nr. 71 John-ren Chen: Energieversorgung, Industrialisierung und Entwicklung einer Dualvolkswirtschaft mit außenwirtschaftlichen Beziehungen, September 1978
- Nr. 72 Hans W. Gottinger: Structural Characteristics of Economic Models: A Study in Complexity, December 1978
- Nr. 73 Reinhard Selten: A Note on Evolutionarily Stable Strategies in Asymmetric Animal Conflicts, December 1978
- Nr. 74 Joachim Rosenmüller, Wulf Albers, Axel Ostmann und W.F. Richter: Projekt Standortspiele, Dezember 1978
- Nr. 75 Joachim Rosenmüller: Selection of Values for Non-Sidepayment Games, December 1978
- Nr. 76 John-ren Chen: Verlauf der ökonomischen Entwicklung einer Dualvolkswirtschaft in der Phase der endogenen Lohnbestimmung - mit besonderer Berücksichtigung der Probleme der Energieverteterung und Devisenknappheit, Januar 1979
- Nr. 77 Axel Ostmann: On Location Conflicts and their Fair Solution Concepts, March 1979
- Nr. 78 John-ren Chen: Fertilizer and Development of an Agricultural Economy, March 1979
- Nr. 79 Hans-W. Gottinger: Complexity, Bounded Rationality and Problem-Solving, March 1979
- Nr. 80 John-C. Harsanyi und Reinhard Selten: A General Solution Theory for Cooperation and Conflict, April 1979
- Nr. 81 Wolfram F. Richter: Social Choice for Bliss-Point Problems, May 1979

Nr. 82 Gottinger, H.-W.: The Simple Economics of Screening Programs:
An Application of Decision Analysis to Medical Screening, May 1979