76. R. N. Sawyer, *Environ. Res.* **13**, 146 (1977).
77. *TLV's, Threshold Limit Values for Chemical Substances in Workroom Air Adopted by ACGIH for 1980* (American Conference of Governmental Industrial Hygienists, Cincinnati, Ohio, 1980).
78. *Asbestos—General Information* (Asbestos Information Association of North America, Washington, D.C., 1975).
79. M. Villeco, *Archit. Forum* **133**, 50 (1970).
80. A. N. Rohl and R. N. Sawyer, paper presented at the International Symposium on Indoor Air Pollution, Health, and Energy Conservation, Amherst, Mass., 13 to 16 October 1981.
81. R. L. Murphy, B. W. Levine, F. J. Al Bazzaz, J. J. Lynch, W. A. Burgess, *Am. Rev. Respir. Dis.* **104**, 576 (1971).
82. Occupational Safety and Health Administration, "Standard for exposure to asbestos dust," *Fed. Regist.* **37**, 11318 (1972).
83. "Revised recommended asbestos standard," *U.S. Dept. Health Educ. Welfare Publ. (NIOSH) 77-169* (1977).
84. R. N. Sawyer and C. M. Spooner, *U.S. Environ. Prot. Agency Publ. EPA-450/2-78-014* (1978).
85. U.S. Environmental Protection Agency, "Amendments to asbestos standard," *Fed. Regist.* **43**, 26372 (1978).
86. Workshop on Indoor Air Quality Research Needs, Interagency Research Group on Indoor Air Quality, Leesburg, Va., 3 to 5 December 1980.
87. "Position statement on indoor quality," *J. Air Pollut. Control Assoc.* **32**, 548 (1982).
88. The American Society of Heating, Refrigeration, and Air Conditioning Engineers, Inc., "Position statement on indoor air quality," approved by the ASHRAE board of directors on 1 July 1981.
89. J. L. Repace, *Environ. Int.* **8**, 21 (1982).
90. K. Sexton and R. Repetto, *ibid.*, p. 5.
91. This work was supported in part by NIEHS grant ES-01108, EPRI grant RP-1001, and EPA grants 68-02-3201 and 2974. We thank the members of the staff and faculty at Harvard University who commented on earlier versions of this manuscript: M. J. Ellenbecker, J. S. Evans, M. W. First, R. Letz, D. W. Moeller, P. B. Ryan, S. R. Thomas, R. D. Treitman, and W. A. Turner. We are also grateful to A. Maskell who typed the article and provided editorial comments.

Cost-Effective Priorities for Cancer Prevention

Milton C. Weinstein

Summary. Faced with limited resources, the United States must set priorities for research to identify preventable causes of cancer. A quantitative approach to priority setting, based on principles of decision analysis and cost-effectiveness analysis, can offer guidance in this process. An illustrative application of such a model suggests that the National Institutes of Health–supported clinical trial of dietary β-carotene offers a greater expected reduction in cancer mortality per research dollar than carcinogen bioassays of high-volume industrial chemicals such as *p*-dichlorobenzene. National research priorities should reflect the relative cost-effectiveness of such investments.

Environmental factors are responsible for 80 to 90 percent of cancer deaths in the United States (*1, 2*). This conclusion, which once aroused considerable controversy, is now generally accepted, provided that the "environment" is broadly defined to include not only industrial chemicals and pollution, but also diet, reproductive behavior, and other elements of life-style and culture, as well as such natural phenomena as infectious agents and nonionizing radiation. Doll and Peto have placed the contribution to U.S. cancer mortality of occupational and environmental exposures to industrial chemicals at less than 5 percent, including 2 percent due to asbestos (*2*).

Growing hope during the 1970's that cancer could be controlled in large part by detecting and eliminating carcinogens has been tempered during the 1980's by the sober realization that preventing cancer will not be simple. Epidemiologic data firmly support the proposition that most cancers are preventable in principle, but current understanding leaves us far short of being able to prevent most cancer in fact. The challenge of the coming decades will be to identify the specific agents that cause or prevent cancer and, after identifying them, to develop and implement interventions to alter human exposure to them.

The problem of identifying carcinogens in the environment seems formidable enough when attention is focused on the 70,000 or so industrial chemicals in production. The cost of testing this inventory of chemicals, let alone the thousands of new chemicals entering production each year, would be huge. Even if financial cost were not a constraint, the limited supply of toxicologists and laboratories would constrain the volume of long-term bioassays.

Epidemiologic insights should, however, lead us to examine the priority-setting problem in a broader framework. If industrial chemicals other than asbestos account for 3 percent of cancer deaths, the prospect of saving 12,000 lives each year (out of 400,000 cancer deaths) should encourage us to discover the key industrial agents and control exposure to them. But if 35 percent of cancer deaths are related to diet (*2*), efforts to discover dietary factors in cancer might deserve an even greater claim on resources.

Toxicologic studies of industrial chemicals and epidemiologic studies of dietary agents are, in general, funded from different budgets, and might seem not to be in competition for the same limited resources. For the society as a whole, however, it is imperative to ask how best to spend resources in the general domain of cancer prevention. Priorities need to be set among alternative research strategies for detecting carcinogenic and anticarcinogenic agents, and such priority setting should encompass the full range of environmental factors (broadly defined) in cancer prevention.

This article illustrates a quantitative approach to priority setting, based on principles of cost-effectiveness and decision analysis. It also shows how the approach may be used to compare the cost-effectiveness of toxicologic studies of industrial chemicals and prospective trials of dietary constituents. The industrial chemical examined is *p*-dichlorobenzene, the active ingredient in mothballs. The cost-effectiveness of a randomized prospective trial of dietary β-carotene, a close relative of vitamin A, is also assessed. This comparison and other considerations lead to policy implications regarding the optimal use of resources in investigating the cancer-related effects of environmental agents.

Uncertainty is inherent in this kind of prospective analysis, and the attempt to quantitate this uncertainty may make some readers uncomfortable. However, policy decisions must and will be made in the face of uncertainty, and analysis

The author is professor of Policy and Decision Sciences, Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115.

can help to organize information, expose the sources and magnitude of uncertainty, and facilitate rational debate about policy options. Further, if major changes in estimates do not change the main conclusions, confidence in those conclusions is increased despite the uncertainty.

An Analytic Framework for Priority Setting

Health effectiveness. Let us assume that the most important objective of laboratory or clinical studies of the carcinogenic effects of specific agents is to prevent cancer mortality, and that the ultimate value of such studies can be measured by the number of cancer deaths prevented or the number of years of life saved (*3*). Overall, the number of years of life saved may be approximated by multiplying the number of cancer deaths averted by 12 years, the mean loss in life expectancy per cancer death in the United States (*4*). An advantage of using life years saved as the measure of effectiveness is that it is general enough to permit comparisons with uses of resources in widely different health programs.

At least three steps are needed to translate a carcinogen bioassay or epidemiologic study into reduced cancer mortality. The test system must detect an effect, the study result must lead to intervention to alter exposure to the substance in question, and the intervention must lead to reduced cancer mortality. The last step implies that the observed effect must be real and not an artifact of the experimental or statistical methods used.

These events are not guaranteed to occur; they are probabilistic in that we do not know in advance whether each of them will occur (for example, whether the test will uncover an effect). Thus, probabilistic reasoning is needed to predict the benefits of studies; a priori, we must settle for a statistical measure of the expected, or average, number of cancer deaths averted or life years gained. Both objective and subjective probabilities are involved.

Decision analysis provides a framework for calculating the expected values of alternative testing strategies (*5*) from estimates of the quantities listed in Table 1.

Table 1. Some factors influencing the expected value, or effectiveness, of a cancer study.

1. Prior probability of effect on cancer mortality and its magnitude
   - Previous tests (in vitro, in vivo)
   - Epidemiologic evidence
   - Biologic understanding or theory
2. Sensitivity of the test system
   - Sample size
   - Technical design features
   - Relevance to human exposures
   - Tolerance for false positives (positivity criterion)
3. Effect of research findings on behavior and exposure
   - Regulation
   - Self-regulation by industry
   - Personal behavior
4. Cancers prevented, over time, given effect and exposure change
   - Potency or relative risk
   - Exposure change
   - Latency
   - Lag to implementation

The prior probability of effect is the best estimate, prior to the study, of how likely it is that the agent in question affects human cancer mortality. For example, one might assess that there is a 10 percent chance that chemical X is a human carcinogen on the basis of previous tests (in vitro and in vivo), epidemiologic evidence (including negative evidence), and structure-activity considerations.

The sensitivity of the test system is the probability that the test will show a positive result if an effect is truly present. It depends on design elements such as sample size and dose of agent, as well as on such technical design features as allocation of treatments, quality control, and observer reliability. If one is using a nonhuman test system to detect effects in humans, another source of insensitivity (false negatives) arises from interspecies differences in response. Finally, the sensitivity depends on the methods of statistical analysis and on the threshold chosen for calling a result positive.

The effect of a study result on behavior and exposure may depend on the actions of many parties. Government may regulate; industry may self-regulate; individuals may alter personal behaviors. It is difficult to predict what actions would follow from specific test outcomes, or even to assign probabilities to these actions. But to fail to assess the prospects for effective intervention would implicitly assign equal value to study results that would probably lead to intervention and results that would only raise anxiety levels or elicit reactions of denial.

The number and timing of cancer deaths prevented may be calculated from estimates of five items: an exposure-response relation, summarized perhaps by a measure such as carcinogenic potency or relative risk; magnitude of exposure change; latency of biological effect (w); lag between the decision to undertake the study and its findings (L_1); and lag between study findings and altered exposure (L_2). The result will typically be a stream of benefits over time, beginning at the end of the lag-plus-latency periods and extending into the indefinite future (Fig. 1). A linear exposure-response relation is assumed over the range of actual exposures, with potency k defined as the number of cancer deaths prevented per year per unit of exposure change. The model can easily be generalized to allow for nonlinear exposure-response relations.

The expected, or average, value of a study may be expressed mathematically as follows. Let us assume a discrete formulation with I possible values of potency, k_i ($0, k_1, k_2, \ldots, k_{I-1}$), each assigned a prior probability, p_i = probability $[k_i]$. Thus p_0 may refer to the prior probability of no effect; p_1 to the prior probability that the potency is k_1 cancers prevented per year per unit of exposure reduction, and so forth. Next, let r_j represent each of J possible study results ($j = 0, \ldots, J-1$); we assess the probability of result r_j given true effect k_i as q_{ij} = probability $[r_j \mid k_i]$. Then, let e_m represent each of M possible changes in exposure ($0, e_1, \ldots, e_{m-1}$), resulting from regulatory control or behavior change, and we assess the probability of exposure change e_m given result r_j (and all prior information) as z_{jm} = probability $[e_m \mid r_j]$. Then the annual expected benefit (or effectiveness, E) of the study is given by

$$E = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \sum_{m=0}^{M-1} (z_{jm} q_{ij} p_i) k_i e_m \qquad (1)$$

Cost. Testing costs include costs of protocol design, implementation, data collection, and analysis. Strictly speaking, however, it would be wrong to balance only these costs against any expected benefits. Interventions to alter exposure are also costly, at least as perceived ex ante; otherwise, why not reduce all exposures to suspected carcinogens to zero? The main analysis excludes the cost of intervention from consideration. Because this would be inappropriate if resources for lifesaving were viewed as constrained in the domain of public health interventions (*6, 7*), costs of intervention are introduced later in the discussion of the examples used to illustrate the model.

The cost-effectiveness ratio. Given estimates of the cost of testing (C) and of the expected number of cancer deaths prevented (E) for a range of studies, the

following question arises: Given that resources do not permit undertaking all studies with positive expected health benefits ($E > 0$), how do we set priorities? The answer, given the objective of maximizing the expected reduction in cancer mortality, centers on the cost-effectiveness ratio, that is, the cost per cancer death averted, for each study being contemplated. If studies are ranked in increasing order of this ratio, and undertaken in the implied order of priority, the total health-effectiveness of the overall testing program will be maximized (*8*).

In forming the cost-effectiveness ratio, one must express costs and benefits in temporally comparable units. Recall that we have an expenditure at time t_0, followed by a delay of $L_1 + L_2 + w$ years before health benefits accrue (Fig. 1). Suppose that the number of cancer deaths averted beginning at time $t_f = t_0 + L_1 + L_2 + w$ is E^* per year. In order to render costs and benefits temporally comparable, we will convert the cost, C, into an equivalent constant annual stream, C^*, commencing at future time t_f, where

$$C^* = rC\,(1 + r)^{t_f - t_0} \qquad (2)$$

and r is the real (that is, inflation-corrected) long-term discount rate, taken to be 0.05 in the following analyses. C^* is analogous to the annual payment on a long-term mortgage at rate r whose principal value is C, but whose first payment is deferred for $t_f - t_0$ years (*9*). The ratio C^*/E^* is then the time-corrected cost-effectiveness ratio (*10*).

Cost-Effectiveness of the Bioassay of *p*-Dichlorobenzene

The carcinogen bioassay in small rodents (CBSR) is the mainstay of carcinogenesis testing in the United States. The National Toxicology Program (NTP) of the federal government has published over 200 reports on chemicals tested under the CBSR, including 27 completed in 1982. Hundreds more chemicals have been tested by private groups, including industry.

Separate from the government's own testing program, the Environmental Protection Agency (EPA), under authority of Section 4 of the Toxic Substances Control Act (TSCA), may require private industry to test chemicals for carcinogenicity. Substantial effort has gone into setting priorities for testing (*11*). This article examines the cost-effectiveness of testing a chemical that emerged at the top of the priority-setting process that led the EPA to identify *p*-dichlorobenzene as one of the first three chemicals to be tested under TSCA.

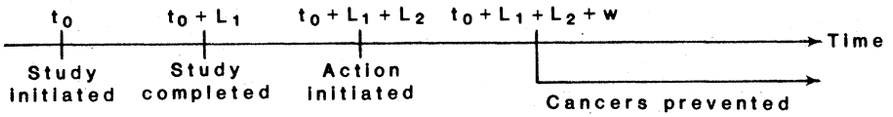

Fig. 1. Timeline for events following initiation of a study. The total delay between initiation of the study and the beginning of the benefit stream is $L_1 + L_2 + w$ (where L_1 is the lag to end of study, L_2 is the lag to initiation of intervention, and w is the cancer latency period).

Approximately 33 million kilograms of *p*-dichlorobenzene were produced in the United States in 1978, of which an estimated 24 million kilograms were released into the air. Of the industrial output, 55 percent is used in space deodorants, 35 percent in mothballs, and 10 percent in a variety of products including pesticides, dyes, floor waxes and finishes, abrasives, and agricultural chemicals (*12*). Despite its widespread use, this compound has not been adequately tested for carcinogenicity.

Prior probabilities and potency. No fully satisfactory basis exists for estimating the probability that a particular chemical is a human carcinogen. Historical experience with the CBSR is a misleading guide because selection of chemicals by the NTP often is based on scientific reasons related to molecular structure rather than public health considerations such as extent of exposure and environmental persistence (*13*). In one random survey of compounds tested prior to 1974 it was estimated that 5 percent of compounds were carcinogenic (*14*).

Information from short-term tests would raise or lower the probability for any particular chemical, depending upon the results (*7*). Apparently, *p*-dichlorobenzene is not mutagenic in the Ames test (*15*). All things considered, let us take 10 percent as the prior probability that *p*-dichlorobenzene is a carcinogen.

This still leaves open the question of how potent it might be. For simplicity, let us apply a point estimate of 0.005 lifetime cancers per milligram per kilogram of body weight per day. This is based on the data compiled by Crouch and Wilson on carcinogenic potency of chemicals tested in animals and man (*16*). This potency is treated as if it applied to fatal cancers only.

Sensitivity of test system. Of 26 known human carcinogens, 18 have been subjected to an adequate CBSR; of these, only two (arsenic and benzene) are not rodent carcinogens (*17*). Thus, one has a rough estimate that 16/18 (89 percent) of human carcinogens are rodent carcinogens. Assuming a statistical power of 0.9, the overall test sensitivity is $(0.89)(0.9) = 0.80$.

Exposure reduction. Occupational exposures are estimated from EPA data to average 28 mg/kg-day for 5000 workers (*18*), and environmental exposures from the air are estimated to average 0.7 μg/kg-day for each of 230 million Americans (*19*).

It is impossible to predict what the response of government, industry, or consumers would be if mothballs were found to cause cancer. Let us assume an expected 50 percent reduction in exposure levels given such a finding (*20*).

Timing of cancers prevented. The CBSR typically requires at least 6 years (L_1) including planning, analysis, and reporting. Allowing at least another 4 years (L_2) for public response, and a carcinogenic latency period of 20 years (w), the total delay would be 30 years before cancer deaths were actually prevented.

Cost of the bioassay. In fiscal year 1981, the NTP let 16 private contracts for CBSR's on 49 chemicals. The mean contract budget per chemical tested by feeding was $467,000. This calculation assumes a cost of $500,000.

Cost-effectiveness calculation. Total occupational and environmental exposure to this compound (e_1 in Eq. 1) is calculated from the above data to be 3×10^5 person-mg/kg-day. Multiplying by the assumed potency, k_1 (0.005 lifetime cancer deaths per person-mg/kg-day divided by 70 years per lifetime), yields 21 cancer deaths per year potentially averted.

The expected effectiveness of testing equals this potential benefit, times the prior probability of carcinogenicity ($p_1 = 0.1$), times the test sensitivity ($q_{11} = 0.8$), times the expected exposure reduction ($z_{11} = 0.5$), so that $E^* = 0.85$ cancer death averted per year.

From Eq. 2, we calculate the annualized cost of the bioassay, at the point of benefit, as (0.05) ($500,000) $(1.05)^{30}$ = $110,000. The cost-effectiveness ratio, then, is (C^*/E^*) = ($110,000 per year)/(0.85 cancer per year) = $130,000 per cancer prevented, or about $11,000 per year of life saved. This excludes the social and economic cost of actually reducing exposure to *p*-dichlorobenzene.

Cost-Effectiveness of the Prospective Trial of β-Carotene

Let us now examine the cost-effectiveness of prospective studies of dietary factors in human cancer. Specifically, let us consider the ongoing prospective trial of dietary β-carotene.

Epidemiologic evidence suggests that retinoids and carotenoids (that is, vitamin A) may contribute to cancer prevention. Peto *et al.* cite 20 dietary studies, of which ten found relative risks of 1.5 to 3.0 for low versus high vitamin A diets, seven found relative risks of 1.3 to 1.5, and only three found no statistically significant effect (*21*). A more recent study of the correlation between cancer incidence and β-carotene consumption (rather than total vitamin A or consumption of vegetables) found a relative risk for lung cancer of 7.0 for all subjects and 8.1 for smokers only (*22*). In addition, several studies in mice have shown that β-carotene can reduce or delay tumor incidence (*21*).

The epidemiologic results are not conclusive for several reasons. One reason is that recall of dietary data is imperfect; another is that the observed negative association between β-carotene consumption and cancer might be an artifact stemming from the carcinogenic effect of animal fat and an inverse association between β-carotene and animal fat levels in the diet. In at least three epidemiologic studies (*23, 24*), however, no association was found between fat or fiber consumption and the cancer under study (larynx, lung, or breast), while a strong protective association with vitamin A consumption was found.

Only a prospective study, in which β-carotene is administered independently of specific foods, can resolve this question. A double-blind controlled trial of β-carotene (30 milligrams every 2 days) in American physicians was funded by the National Institutes of Health, after some deliberation, as an add-on to a trial examining the relation between aspirin and myocardial infarction, using a 2 by 2 factorial design. The protocol calls for a 5-year intervention and follow-up (*25*).

The following data and assumptions were used in the cost-effectiveness analysis of the β-carotene trial.

Prior probabilities and potency. Estimates of association between β-carotene and cancers of the lung, bladder, larynx, esophagus, and breast were derived from epidemiologic studies. Two studies provided estimates of relative risk for lung cancer. Mettlin *et al.* (*24*) found a 70 percent increased risk of lung cancer in that half of a population with a lower β-carotene consumption compared to the upper quartile of the population. In a similar study of β-carotene consumption, Shekelle *et al.* (*22*) reported a sevenfold risk for males in the lowest quartile compared to the highest quartile. In this article it assumed that an increase in intake to 15 mg/day would correspond to moving to the low-risk quartile, and the lower of the risk estimates from the two studies is used. A total of 30,600 lung cancer deaths would therefore be averted per year, or 32 percent of all such deaths in the United States.

Estimates for cancers of the bladder, larynx, esophagus, and breast were calculated analogously from epidemiologic data (*26*) as 3300, 1500, 2400, and 6900 deaths averted per year, respectively, or 35, 50, 30, and 20 percent of cancer deaths at these sites. With a 5 percent reduction in cancer mortality being assumed at all other sites combined, an additional 13,000 deaths would be averted, bringing the total to 58,000 cancer deaths.

Therefore, the potential reduction in cancer mortality if the β-carotene hypothesis were true and if the population altered its dietary habits, is estimated to be 60,000 deaths per year, or a 15 percent reduction.

Finally, we need a subjective estimate of the probability that the hypothesis is, in fact, correct. Let us use a subjective probability estimate of 10 percent.

Test system sensitivity. The statistical power of the study depends on at least four factors: the magnitude of any true effect, the sample size, the duration of the study, and the latency period prior to manifestation of the effect. Approximately 15 percent of all male U.S. physicians aged 50 to 75 years have been enrolled in the study. The approximately 20,000 subjects are divided randomly between treated persons and placebo controls, and follow-up will be for 5 years.

Assuming a 2-year latency period, and a 15 percent reduction in male cancer incidence during years 3 through 5, we would expect 3.17 percent cancer incidence in the controls and 2.69 percent incidence in the treated group. With a significance level of .05, under these assumptions, the probability of the study detecting the effect, if present, would be 64 percent (*27*). Since there is no issue of interspecies correlation, 0.64 is used as the estimate of test system sensitivity.

Exposure change. The public health impact of a positive finding would depend on the responses of both public health officials and private individuals. Public health officials could declare such a finding cause for a major public health campaign analogous to fluoridation of public water supplies. Already, dairy products are fortified with vitamins; they could be fortified with higher doses of β-carotene, provided the public would accept foods with a slight orange tint. Failing such a mass intervention, individuals could be urged to increase their β-carotene intake by promotional campaigns; or subsidies for high β-carotene foods could be increased. Finally, individuals may elect to take β-carotene as a drug; several commercial preparations are available.

What proportion of the population would modify their diet in response to a positive finding? What is the likelihood of a public health initiative to fortify foods? These are difficult questions, but it might be easier for parents to inculcate a tolerance for carrots than an abhorrence of smoking. Let us suppose, subjectively and perhaps conservatively, that 10 percent of the potential benefit would be realized by some combination of public and private initiatives.

Timing of cancers prevented. Let us assume a lag equal to the study duration (5 years), plus an additional 10 years for dissemination and latency.

Cost of the study. The budget for the combined study of β-carotene and cancer and of aspirin and myocardial infarction is $4,000,000. Let us attribute the full cost to the β-carotene study, recognizing that this tends to overestimate its true incremental cost.

Cost-effectiveness calculation. Under our central assumptions (64 percent study power, 10 percent compliance, 15 percent cancer mortality reduction), the estimated reduction in annual cancer mortality if an effect is present would be 3840. Multiplying by the prior probability of 0.1, the expected annual benefit from the study is 384 cancer deaths averted.

From Eq. 2, with $r = .05$ and $t_f - t_0 = 15$ years, we calculate the annualized cost of the study to be $420,000 per year. The cost-effectiveness ratio is C^*/E^* = ($420,000 per year)/(384 cancer deaths per year) = $1100 per cancer death prevented, or about $91 per year of life saved. This is approximately 1 percent of the corresponding estimate for the bioassay of *p*-dichlorobenzene.

Comparative Cost-Effectiveness and Sensitivity Analysis

To summarize, the expected cost per year of life saved is expected to be $11,000 for a rodent bioassay of *p*-dichlorobenzene and $91 for a prospective trial of β-carotene. Both figures are for research studies to establish harmful or

beneficial effects, and exclude the costs of intervention.

The data and assumptions underlying these calculations are soft. But is the 100-fold difference in cost-effectiveness large enough to withstand even rather large errors in the estimates? Let us approach this question by means of sensitivity analysis.

In the calculation of cost-effectiveness for the bioassay of *p*-dichlorobenzene, the six critical parameters were as follows: the exposure estimates for workers and for the general public, the assumed carcinogenic potency, the prior probability of carcinogenicity, the sensitivity of the test system, the proportion reduction in exposure, and the lag and latency periods.

The 5000 workers assumed to be exposed to two-thirds of the maximum allowable time-weighted average concentration of the potential carcinogen include those directly involved in the manufacturing processes. We might have assumed another 50,000 to be exposed to 2 mg/m^3, the equivalent of a mothball-filled closet, but this would add only 100,000 person-mg/m^3 to the original occupational estimate of 1,500,000 person-mg/m^3. Even if all 500,000 persons employed in the industry were exposed to 2 mg/m^3, this would add only 60 percent to the occupational exposure estimate, or about 30 percent to the overall exposure estimate.

The assumed carcinogenic potency is already five times the human potency of benzene, and more than double the mouse potency of ethylene dichloride (*28*). Moreover, we are assuming all of the inhaled chemical to be absorbed and are using a conservative linear dose-response model. Although the estimate of potency may be conservatively high already, we double it to 0.01 cancer deaths per person-mg/kg-day in the sensitivity analysis. The prior probability of 10 percent is already as high as is consistent with the negative evidence from short-term tests.

The sensitivity of the test system (assumed to be 80 percent) exerts little leverage on the analysis, and is as high as is reasonable given statistical and interspecies considerations. Similarly, the assumed 50 percent reduction in exposure seems as high as is realistic, although conceivably a virtual 100 percent reduction could be achieved if the chemical were banned. Finally, the latency period could be shorter than 20 years; as an extreme case for sensitivity analysis, we take it to be zero.

Under all of these extreme assumptions, the cost-effectiveness ratio for the bioassay of *p*-dichlorobenzene would fall from $11,000 per year of life saved to approximately $1,000.

On the other side, it is unlikely that we have been overly optimistic in our estimates about β-carotene. If we assumed a 10 percent (rather than 15 percent) reduction in cancer mortality, with prior probability 10 percent, the cost-effectiveness ratio would change from $91 to $240 per year of life saved, still extraordinary by most standards. Thus, the highest plausible figure for the β-carotene study is still well below the lowest plausible figure for a CBSR assay of *p*-dichlorobenzene. Plausible assumptions in the opposite directions would have increased the divergence from two orders of magnitude to three or even four orders of magnitude.

One caveat in advocating studies such as the β-carotene trial is that the possibility of a false-negative finding may be seen as unacceptably high. A negative result may engender future public mistrust of public health information and may make further research on diet and cancer more difficult to justify. It will be difficult to explain that, even if there were a 15 percent reduction in cancer, the study had a 36 percent chance of missing it. One remedy would be to increase the duration of the study. Under our previous assumptions, a 10-year study would increase the statistical power from 0.64 to approximately 0.95, while perhaps doubling the cost. Thus, the cost-effectiveness ratio would increase somewhat (though still less than $200 per year of life saved), but the chances of a falsely negative result would be reduced considerably.

Considering the Cost of Intervention

Let us now consider the cost of intervention to alter exposures to these substances.

The social and economic cost of banning major uses of *p*-dichlorobenzene would be great. The gross annual primary sales of this compound are approximately $30 million. Using this figure as an estimate of the economic benefits forgone if a ban were implemented, taking our estimate of 21 cancer deaths averted per year if the chemical is a carcinogen, and assuming that there is at least a 5 percent chance of a falsely positive bioassay result, we obtain a cost-effectiveness ratio of $2.2 million per life saved, or $186,000 per year of life saved (*29*). Perhaps, on the other hand, acceptable and economical substitute products could be found, or minor adjustments in work practice and prudent use in the home could reduce exposures considerably at a minimal cost. In any case some economic dislocations would surely be felt.

The annual retail cost of taking 15 mg of β-carotene daily is $36.50 (*30*). The long-run cost, with generically available β-carotene, might be in the $20 to $30 range. Less costly would be fortification of foods such as milk or butter, or individual diet modification; two 100-g servings of carrots contain 13.4 mg of β-carotene. The side effects of β-carotene consumption are benign (*31*). Most prominent is a coloring of the skin which some people find appealing.

Suppose the cost of an annual regimen of β-carotene were $30. If 200,000,000 Americans paid this price (excluding small children), the annual bill would be $6 billion. If this could prevent 60,000 cancers a year, and if the ratio of the true-positive to false-positive study results is as assumed previously, the steady-state cost per cancer death averted would be $170,000, or about $14,000 per year of life saved. The cost-effectiveness ratio would be lower if the program were targeted at older age groups. Even as an upper bound, this cost per year of life saved compares favorably to many preventive medical interventions in common use, such as treatment of high blood pressure (*32*) and cancer screening (*33*). It is one or two orders of magnitude lower than ratios estimated for occupational and environmental health measures aimed at cancer prevention (*34*). It is also likely to be substantially lower than the corresponding estimate for a ban on *p*-dichlorobenzene, which reinforces the conclusion from the main analysis in which the costs of intervention were excluded.

Policy Implications

The β-carotene trial appears to be an excellent investment in health resources. The carcinogen bioassay of *p*-dichlorobenzene may also be a reasonable use of resources, although not nearly as high a priority as the β-carotene trial.

Can it be concluded more generally that, as an approach to cancer prevention, studies designed to test dietary hypotheses in humans are likely to be more productive (in terms of health benefits per dollar spent) than carcinogen bioassays in small rodents? The answer depends on the degree to which the examples chosen for analysis are typical of their classes. *p*-Dichlorobenzene was given the highest priority for testing by the EPA on the basis of human exposure and other considerations, and may therefore be considered a best case. β-Caro-

tene may also be considered a best case, since it is one of relatively few dietary factors that are now ready for prospective study. But even if the prior estimates for other agents such as vitamin E, vitamin C, and selenium are ten times less favorable than that for β-carotene, such studies appear to be well worth the costs when the expected benefits are compared to those that might be derived from animal bioassays of industrial chemicals or to current uses of health care resources in preventive and curative medicine.

As an immediate policy implication, it may be concluded that, apart from purely scientific considerations, industrial chemicals should be carefully screened on the basis of exceptionally high exposure or strong prior evidence of carcinogenic potential prior to the initiation of long-term studies. The chemical-by-chemical approach to discovering carcinogens appears to be less cost-effective than other uses of the same resources.

Large-scale national (and perhaps international) tests of dietary hypotheses seem to be promising and cost-effective uses of health resources. This analysis underscores the value of basic and epidemiologic research to identify new dietary hypotheses. Given the strong evidence linking diet to cancer, it seems likely that large-scale epidemiologic investigations will generate hypotheses as promising as β-carotene, and some may have major implications for cancer mortality in the United States.

Several countervailing observations might seem to lessen the strength of these conclusions. First, the population of potential subjects for studies like that of β-carotene is limited. It would be difficult to mobilize enough subjects to conduct more than a few such trials at one time.

Second, changes in personal behavior may be difficult to effect even if dietary factors are found to be protective. The response to data on smoking and lung cancer has been smaller and slower than might have been hoped for. However, a large share of the decline in cardiovascular mortality may reasonably be attributed to changes in health behaviors such as blood pressure control, diet, as well as tobacco use (*35*).

Psychological and political factors may also favor continued vigilance over industrial chemicals, despite unfavorable cost-effectiveness ratios, and such factors must be recognized as legitimate. Involuntary exposures to cancer risks may be feared more than voluntary exposures, and public officials do not gain much politically by helping their constituents discover what foods are good or bad for them.

As a final caveat, this framework assumes that the value of the information yielded by a study lies in its ability to influence decision-makers—in industry, in government, and private individuals—to alter exposures to the agent in question. To the degree that a study contributes to scientific knowledge per se, however, it may lead indirectly to future improvements in public health, and such considerations ought to affect the estimated value of a study. From this perspective, the value of basic research on mechanisms in carcinogenesis should not be underestimated as a result of myopic applications of policy models such as the one proposed here.

This article began with the premise that the criterion for health resource allocation ought to be health benefit, somehow defined. This led to the criterion of cost-effectiveness. But how can the institutions of our government—and society, more generally—be structured to make these tradeoffs in domains as diverse as animal toxicology and human epidemiology, or cancer prevention and cancer treatment? The mission of the National Cancer Institute would seem to require that it reexamine these priorities and allocate resources accordingly. Moreover, the debate about hospital cost containment should not be carried on in isolation from concerns for environmental health and chemoprevention. The common value of all of these programs to society is health. As resources for health-related activities become increasingly constrained, we must ask ourselves anew the question of whether we are spending our resources wisely. Failure to do so may result in lost opportunities to control the most dread diseases of our society.

References and Notes

1. E. L. Wynder and G. B. Gori, *J. Natl. Cancer Inst.* **58**, 825 (1977); J. Higginson and C. S. Muir, *ibid.* **63**, 1291 (1979).
2. R. Doll and R. Peto, *ibid.* **66**, 1192 (1981).
3. Other, more complicated measures could incorporate concerns for the adverse effects of cancer on the quality of life, or the economic burden of cancer. It is unlikely, however, that such additional considerations would affect the relative cost-effectiveness of alternative cancer research investments.
4. J. Cairns, *Cancer: Science and Society* (Freeman, San Francisco, 1978). This figure may be adjusted for specific tumor sites or demographic groups, but such adjustments will rarely make much difference for cancers in the adult age range.
5. The measure of effectiveness used here is conceptually identical to the "expected value of sample information" of statistical decision theory [H. Raiffa, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty* (Addison-Wesley, Reading, Mass., 1968)].
6. One remedy might add the cost of intervention (weighted by the probability that the study result will lead to intervention) to the cost of performing the study to give the total expected cost to society. Alternatively, one might weight the cost of intervention differently from the testing costs, where the weights reflect the relative lifesaving potential (that is, opportunity costs) of resources diverted from testing or intervention, respectively (*7*). The problems with ignoring costs of intervention are brought to the fore by the observation that the costs of false positives—that is, agents for which intervention is inappropriately implemented as a consequence of a positive study result despite the absence of a true effect—will be excluded from the formal analysis.
7. M. C. Weinstein, *Public Policy* **27**, 333 (1979).
8. The rationale and mechanics of the cost-effectiveness approach to health program evaluation are discussed in M. C. Weinstein and W. B. Stason, *N. Engl. J. Med.* **296**, 716 (1977); M. S. Thompson, *Benefit-Cost Analysis for Program Evaluation* (Sage, Beverly Hills, Calif., 1980); and K. E. Warner and B. R. Luce, *Cost-Benefit and Cost-Effectiveness Analysis in Health Care: Principles, Practice, and Potential* (Health Administration Press, Ann Arbor, Mich., 1982). Strictly speaking, the optimality of the cost-effectiveness ratio rule depends on the assumption that investments (for example, studies) are mutually independent and divisible; in practice, deviations from these assumptions rarely alter the priorities implied by the ranking by cost-effectiveness ratios, and are unlikely to affect the examples considered here.
9. The amortization formula presumes an infinite horizon. If the benefits of altered exposure are expected to be of finite duration, then the appropriate C^* would be slightly greater.
10. An alternative to this method of temporal adjustment would be to calculate the present value of all costs and benefits at the discount rate r. This procedure, while mathematically equivalent to that chosen here, has two drawbacks. First, it is more convenient, and intuitively appealing, in this context to work with constant streams than with lump sums. Second, the procedure of taking the future amortized values of economic costs may appear to be ethically more reasonable than discounting future lives saved. The two procedures are equivalent, however, in that they increase the relative burden of early costs to their proper value at the time at which benefits are realized.
11. For a review of such priority-setting schemes, see National Academy of Sciences, *Strategies to Determine Needs and Priorities for Toxicity Testing* (National Academy Press, Washington, D.C., 1981), vol. 1.
12. U.S. Environmental Protection Agency, *An Exposure and Risk Assessment for Dichlorobenzenes. Final Draft Report* (Environmental Protection Agency, Washington, D.C., 1981).
13. Although 111 of the first 227 chemicals tested by NTP were carcinogenic in at least one species, one cannot conclude that 50 percent of chemicals are rodent, or human, carcinogens. In fact, since 1979, the proportion testing positive has been falling as the selection of chemicals has become based more on public health concerns.
14. R. L. Dehn and C. T. Helmes, *An Automatic Procedure for Assessing Possible Carcinogenic Activity of Chemicals Prior to Testing* [Report prepared for the National Cancer Institute (Stanford Research Institute, Menlo Park, Calif., 1974)]. The National Academy of Sciences (*11*) is implementing an empirical study in which existing data on carcinogenicity of randomly selected chemicals will be examined, thus providing an updated basis for estimating a prior probability for chemicals in general.
15. T. Lawlor, S. R. Haworth, P. Voytek, *Environ. Mutagen.* **1**, 143 (1979).
16. E. Crouch and R. Wilson [*J. Toxicol. Environ. Health* **5**, 1079 (1979)] fitted data from NTP bioassays in the B6C3F1 mouse and the Fischer rat to a single-hit dose response model with correction for background tumors. Excluding the chemicals with the two most extreme estimated values of b for both species, and averaging the rest of the estimated potencies, a mean potency of 0.002 cancer (lifetime) per milligram per kilogram per day ingested was calculated for B6C3F1 mouse and 0.008 for the Fischer rat. In accordance with the result of Crouch and Wilson that potencies in rodents and men correlate well on average, with a regression coefficient of 1, a point estimate of 0.005 lifetime cancers per milligram per kilogram per day is applied. A model in which potency is measured relative to fractions of maximum tolerated dose, rather than absolute quantity ingested, might lead to more stable and more valid potency estimates, but the data have not yet been analyzed in that form.
17. L. Tomatis, *Annu. Rev. Pharmacol. Toxicol.* **19**, 511 (1979). Benzene is added to the 17 human

carcinogens considered by Tomatis to have been adequately tested.

18. According to the EPA (*12*), estimates of the number of workers exposed range from 500,000 (estimate from the National Occupational Hazard Survey) to fewer than 5000 (an industry estimate prepared by Hull and Co. for the Synthetic Organic Chemical Manufacturing Association). The latter is probably more appropriate at levels of exposure close to the permissible 8-hour time-weighted average (TWA) ambient levels of 450 mg/m^3, or to the empirically observed TWA levels of 300 mg/m^3. Thus it is assumed that 5000 workers are each exposed to 300 mg/m^3 for 8 hours per day, 250 days per year. Assuming each worker inhales 1.2 m^3/hour and absorbs all of the *p*-dichlorobenzene, annual exposure would be (250 days per year)(8 hours per day)(1.2 m^3/hour)(300 mg/m^3) = 720 g/year. Converting to an average daily dose in milligrams per kilogram per day, one obtains (720 g per year)/(70 kg)(365 days per year) = 28 mg/kg-day for 5000 workers.
19. This estimate is based on the EPA's estimated frequency distribution of persons exposed to various ambient levels from 100 μg/m^3 to ≤ 0.1 μg/m^3, and a mean inhalation rate of 1.2 m^3/hour while awake for 16 hours a day, and 0.4 m^3/hour while asleep for 8 hours a day. Environmental exposures from water were calculated on the basis of EPA estimates, and were found to be negligible compared to the air (less than 0.1 percent of the amount inhaled).
20. A report prepared by the NTP for the Senate Appropriations Committee found that of 98 chemicals testing positive in at least one species prior to September 1979, 55 had proposed or final regulations on the record. However, many of these regulations were guidelines and not mandatory standards, and, with the exception of tris-BP (the flame retardant once used in children's sleepwear), it is impossible to document any change in exposure. For example, no new standards were promulgated after it was shown by means of a CBSR that ethylene dichloride is a carcinogen, but there may have been intensified voluntary efforts to reduce exposure to this high-volume industrial chemical.
21. R. Peto, R. Doll, J. D. Buckley, M. B. Sporn, *Nature (London)* **290**, 201 (1981).
22. R. B. Shekelle, S. Liu, W. J. Raynor, Jr., M. Lepper, C. Maliza, A. H. Rossof, *Lancet* **1981-II**, 1185 (1981).
23. S. Graham, C. Mettlin, J. Marshall, R. Priore, T. Rzepka, D. Shedd, *Am. J. Epidemiol.* **113**, 675 (1982).
24. C. Mettlin, S. Graham, M. Swanson, *J. Natl. Cancer Inst.* **62**, 1435 (1979); S. Graham, J. Marshall, C. Mettlin, T. Rzepka, T. Nemoto, T. Byers, *ibid.* **116**, 68 (1982).
25. A 5-year trial may seem short in view of the usual latency periods in carcinogenesis. However, Peto *et al.* (*21*) interpret the evidence from animal studies as suggesting that β-carotene inhibits a late stage in carcinogenesis; if so, its effects would be seen in just a few years.
26. C. Mettlin and S. Graham, *Am. J. Epidemiol.* **110**, 255 (1979); S. Graham *et al.*, *ibid.* **113**, 675 (1981); C. Mettlin, S. Graham, R. Priore, J. Marshall, M. Swanson, *Nutr. Cancer* **2**, 143 (1981); S. Graham *et al.*, *Am. J. Epidemiol.* **116**, 68 (1982). The percent reduction in breast cancer was based on a 3:1 weighting of the relative risks for persons over 55 and under 55 years of age, reflecting the age distribution of breast cancer mortality in the United States.
27. The power estimate, β, is based on a one-tailed test and is given by

$$\beta = \Phi\left\{\frac{N^{1/2}(P_1 - P_0)}{[P_1(1 - P_1) + P_0(1 - P_0)]^{1/2}} - 1.645\right\}$$

where N = sample size per group, P_0 = placebo mortality rate, P_1 = treated mortality rate, and Φ = cumulative Gaussian distribution function. With a two-tailed test, the power estimate would be .52.
28. E. Crouch and R. Wilson, *J. Toxicol. Environ. Health* **5**, 1095 (1979).
29. With a prior probability of 0.1, a true-positive rate of 0.8, and a false-positive rate of 0.05, the probability of carcinogenicity after a positive study is (0.1)(0.8)/[(0.1)(0.8) + (0.9)(0.05)] = 0.64. Hence, the cost-effectiveness ratio is ($30 × 10^6)/(21)(0.64), or $2.2 × 10^6 per cancer death averted.
30. Based on an average retail price of $20 per hundred at discount pharmacies in the Boston area.
31. β-Carotene may be taken safely in large doses, up to 180 mg/day [U.S. Food and Drug Administration, *Evaluation of the Health Aspects of Carotene (Beta-Carotene) as a Food Ingredient*, prepared by the Federation of American Societies for Experimental Biology (PB80-119837, National Technical Information Service, Springfield, Va., 1979)].
32. M. C. Weinstein and W. B. Stason, *Hypertension: A Policy Perspective* (Harvard Univ. Press, Cambridge, Mass., 1976).
33. D. M. Eddy, *Screening for Cancer: Theory, Analysis and Design* (Prentice-Hall, Englewood Cliffs, N.J., 1980).
34. J. D. Graham and J. W. Vaupel, *Risk Analysis* **1**, 89 (1981).
35. R. Levy, *Annu. Rev. Publ. Health* **2**, 49 (1981).
36. Supported by grants from the Alfred P. Sloan Foundation and the Mobil Foundation. I thank D. Atkins for research assistance and J. C. Bailar III, P. Braun, J. Cairns, M. Thompson, and two anonymous reviewers for suggestions.

RESEARCH ARTICLE

Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*

Welcome Bender, Michael Akam, François Karch
Philip A. Beachy, Mark Peifer, Pierre Spierer
E. B. Lewis, David S. Hogness

The bodies of insects are divided into a series of segments. The segments are formed very early in the development of the embryo, and cells from one segment do not, in general, mix with cells from other segments throughout the rest of development (*1*). In the fruit fly *Drosophila melanogaster*, there are mutations that transform parts of segments or entire segments into the form of other segments. These homeotic mutations define genes that direct cells into different developmental pathways in different segments. The bithorax complex in *Drosophila* is one of the best studied clusters of such genes (*2*); these genes determine the developmental fate of many of the thoracic and abdominal segments of the animal. When the whole bithorax complex is deleted, the animal dies late in embryonic development and shows striking changes in the segmental pattern of the embryonic cuticle. The third segment of the thorax and all eight abdominal segments resemble the normal second thoracic segment (*2*). Thus the second thoracic segment, which gives rise to the pair of wings and the second pair of legs in the adult fly, can be considered the developmental ground state, and the bithorax complex directs the more posterior segments to specialized developmental pathways. Individual recessive mutations within the complex give less extreme segmental transformations than those resulting from deletions of the whole complex. These mutations transform part of a segment or segments into tissue appropriate to a more anterior segment, toward the ground state. There are also dominant mutations, which transform a segment or part of a segment into more posterior structures, away from the ground state (*3*). These dominant mutations seem to upset the regulation of genes within the complex and turn on functions in an inappropriate segment.

A genetic map of the complex is shown in Fig. 1. Most of the recessive mutants and several dominant mutants show no cytologically visible rearrangements in the salivary gland polytene chromosomes, and they can be recombined with each other. The recombination distances between some pairs are shown. The recessive mutations *bx* and *pbx* affect development of the anterior and posterior halves, respectively, of the third thoracic segment. In the abdomen, *bxd*, *iab-2*, *iab-5*, and *iab-8* affect the first, second, fifth, and eighth ab-

W. Bender, F. Karch, and M. Peifer are investigators in the Department of Biological Chemistry at Harvard Medical School, Boston, Massachusetts 02115. M. Akam is an MRC Senior Fellow in the Department of Genetics, Cambridge University, Cambridge CB2 3EH, England. P. A. Beachy is a graduate student and D. S. Hogness is a professor in the Department of Biochemistry at Stanford University School of Medicine, Stanford, California 94305. P. Spierer is an investigator in the Department of Molecular Biology, University of Geneva, 1211 Geneva 4, Switzerland. E. B. Lewis is a professor in the Division of Biology, California Institute of Technology, Pasadena 91125.