March 1997

The Normative Nature of Risk Assessment: Features and Possibilities

Carl F. Cranor

Follow this and additional works at: https://scholars.unh.edu/risk

Part of the Environmental Sciences Commons, Life Sciences Commons, and the Social and Behavioral Sciences Commons

Repository Citation


This Article is brought to you for free and open access by the University of New Hampshire – School of Law at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in RISK: Health, Safety & Environment (1990-2002) by an authorized editor of University of New Hampshire Scholars' Repository. For more information, please contact ellen.phillips@law.unh.edu.
The Normative Nature of Risk Assessment: Features and Possibilities*

Carl F. Cranor**

Introduction

Risk assessment is frequently seen as technical or scientific, which it is, at least in part. One cannot, however, treat it in only this way, because it is permeated with normative presuppositions. Recognizing the normative aspects of risk assessment greatly affects how we use it and what we should expect from it, but, importantly, it also opens up possibilities that might not be foreseen or understood if we merely treat it as a part of core scientific disciplines. Here, I focus on aspects of carcinogen risk assessment,1 identify some of its main normative presuppositions and then suggest some inferences from these points.

I

First, the very idea of a risk assessment is normatively laden. A risk is the probability of an undesirable outcome (when probabilities can be assigned).2 Thus, a risk is the probability of an outcome that will be unfortunate or undesirable from some evaluative point of view. This first conceptual point is relatively unimportant, except as I note later, the object of value and its importance may influence how one does the risk assessment and treats some of the uncertainties involved. If the valued object threatened by the risk is sufficiently important vis-a-vis

* This paper was supported in part by a National Science Foundation Grant (SBR-93107995), An Evaluation of the Desirability of Evidentiary Procedures for Identifying Carcinogens.

** Dr. Cranor is Professor of Philosophy and Associate Dean, College of Humanities and Social Sciences, University of California, Riverside. He holds a B.A. (Mathematics), University of Colorado; Ph. D. (Philosophy), University of California, Los Angeles and M.S.L. (Law), Yale Law School. Email: carl@chss.ucr.edu.

1 I focus on carcinogen risk assessment because more tends to be known, and risk assessment has been more highly developed. This focus should not be taken, however, to indicate that carcinogens are the only concern.

2 Nicholas Rescher, Risk: A Philosophical Introduction to the Theory of Risk Evaluation and Management 5 (1983) (when probabilities cannot be assigned to each aspect of the assessment, one has a risk and uncertainty assessment).
the evidentiary values implicit in the risk assessment, it may effect the conceptions and design of the evidentiary procedures.

Second, even at its best risk assessment is imperfect; mistakes will result. The ideal is a harm assessment in which we would assess the actual harms that would result from using carcinogens in our products or having them as byproducts of our activities. The second best case is a risk assessment which aims to estimate the probabilities of unfortunate or undesirable outcomes from using carcinogens. Since probabilities are involved, the estimates may not be correct. Furthermore, in the present state of knowledge because carcinogen risk assessment is permeated with a number of uncertainties, the third best case is a probability and uncertainty assessment; this further increases the possibility of mistakes. Two kinds of mistakes might result: false negatives, wrongly exonerating a toxic substance as non-toxic, and false positives, wrongly condemning a non-toxic substance as toxic. With perfect knowledge there may be correct answers in the real world, but the chosen procedures do not guarantee these outcomes. These possibilities are illustrated in the following table.

<table>
<thead>
<tr>
<th>Procedure shows no association between exposure and disease</th>
<th>No association between exposure and disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>No mistake</td>
<td>False negative — β</td>
</tr>
<tr>
<td>Procedure shows an association between exposure and disease</td>
<td>False positive — α</td>
</tr>
</tbody>
</table>

While it is desirable to have procedures which result in no mistakes — no false positives and no false negatives \(N_{FN}=0; N_{FP}=0\), where \(N_{FN}\) is the number of false negatives and \(N_{FP}\), false positives — it is rare to have perfect procedures, and risk assessment is far from perfect.

Absent perfect procedures, the alternative is to use processes that minimize resulting mistakes. This might be roughly represented as: \(\min(N_{FN}, N_{FP})\). However, there is a problem with this characterization. One cannot minimize two variables at once. The problem is even worse when, as sometimes occurs in risk assessment, the two mistake variables
are mathematically interdependent (see point four below). Moreover, and more importantly, it is not the total number of mistakes that is critical, but different kinds of mistakes have different consequences and importance depending upon the context. For example, in screening for breast cancer or AIDS, false positives are arguably of lesser concern than false negatives; positive outcomes can be followed up with more sensitive tests, while false negatives will wrongly give a false sense of security which could result in patient deaths. In other circumstances, false positives may be more important. In the non-scientific area of the criminal law, it is much more important to avoid wrongly convicting an innocent person (a legal false positive) than it is to avoid wrongly exonerating a guilty person (a legal false negative).

The aim, then, for a procedure which is imperfect, should be to reduce the total costs of mistakes, i.e., reduce mistakes that are more important from some evaluative point of view. Symbolically, this suggests that the aim should be to minimize the sum of the social costs of mistakes: min[(\(N_{FN} \times SC_{FN}\)) + (\(N_{FP} \times SC_{FP}\))], where \(SC_{FN}\) is the social cost of a false negative and \(SC_{FP}\) is the social cost of a false positive. The term “social costs” includes not only money, but also other matters that might be important to us (e.g. suffering, loss of life, being deprived of products or opportunities). This representation suggests that the aim of a procedure should be to minimize the total costs of all mistakes taking into account the number and importance of each kind of mistake. Thus, this might include designing a risk assessment process in one case which had few false positives and more false negatives and in another case one which had few false negatives, and more false positives, since sometimes it is more important to avoid one kind of mistake and sometimes more important to avoid another.

There is a further normative point which should be recognized and incorporated into the symbolic representation above. While it might be desirable to have an assessment procedure that minimizes the total costs of mistakes, the social, monetary and human costs of trying to implement such a procedure could also be high or low. Thus, in designing procedures, the costs of the procedure itself should be part of costs to be minimized. Symbolically, this would be represented as min[(\(N_{FN} \times CF_{FN}\)) + (\(N_{FP} \times CF_{FP}\)) + \(SC_{T}\)], where \(SC_{T}\) is the social cost.
of testing or the cost of using a procedure to reduce some of the mistakes. The point here is that if a particular procedure becomes too costly, either itself or for the value of its information, even though it better reduces the costs of false positives or false negatives than some other procedures, it may have to be modified or abandoned. For example, in the criminal law we can reduce the number of innocent persons who are convicted by making both search and seizure rules more strict and by making the proof requirements more demanding. Both actions would decrease the number of innocent people who are punished, but there are substantial social costs of doing so. Conversely, we could reduce the number of guilty persons who go free by making search, seizure and trial rules very favorable to the prosecution, but again at substantial social costs. Similarly, carcinogen risk assessment procedures could be designed dramatically to reduce either false positives or false negatives, but the costs of the two different alternatives should be evaluated. Thus, in designing such procedures these normative considerations together with the numbers of the different kinds of mistakes that are possible should be taken into account.

The above points suggest a third important consideration: Since risk assessments are not perfect evidentiary procedures and the resulting aim should be to reduce total social costs, such procedures are importantly normative because of the social costs of different kinds of mistakes. The normative portion of the symbolic representation is an on balance judgment of the importance of avoiding one kind of mistake or the other together with normative pros and cons of the procedures used to reduce different mistakes. Thus, such procedures should be designed or used by practitioners with these normative considerations in mind.3

A fourth major point about normative considerations is that attempting to reduce one kind of mistake will often increase other kinds of mistakes. One can show, for statistical studies using relatively small sample sizes and attempting to detect a relatively subtle effect, that one mathematically cannot keep the number of false positives and false negatives both low.4 One cannot reduce both false negatives and

---

3 Who should make these decisions and how they should be made in a democratic society are important topics that cannot be addressed here.

4 S. D. Walter, Determination of Significant Relative Risks and Optimal Sampling Procedures in Prospective and Retrospective Comparative Studies of Various Sizes, 105 Am. J. Epidemiology 387 (1977) and Carl F. Cranor, Some
false positives simultaneously without greatly enlarging the sample size. When such conflicts occur, one must decide the proper balance between one mistake or another. This should be explicitly faced; it is easy to miss in studies where strong conventions about minimizing false positives are presumed or serve as default assumptions. Moreover, decisions about sample sizes, cost and study feasibility also raise normative issues. The specificity and sensitivity of studies can always be improved by increasing sample sizes, but this increases costs and may decrease feasibility. Thus, in the very conception and design of a study, critical tradeoffs may be made between cost and accuracy.

A fifth, related point is that practices and conventions within a field can inadvertently predispose procedures to certain outcomes. In studies to the extent that typical conventions aim to prevent false positives, other things being equal, they increase false negatives. Also, practices, temptations almost, that are part of scientific disciplines may incline scientists to approach such problems as they approach bench science. This may frustrate other goals, i.e., identifying risks to people or the environment. While such practices might initially appear to be the proper application of scientific methods, their application in particular contexts may frustrate aims that risk assessment is designed to serve. As indicated below, the temptation to do a detailed toxicity analysis of every chemical, case-by-case in a science-intensive way, while serving some aspects of scientific accuracy, may frustrate the timely identification, analysis and regulation of substances that pose risks. In epidemiology as some researchers have noted "One can always invoke unmeasured confounders to explain away observational associations. Thus, actions should not depend on the absence of such explanations, for otherwise action would never be taken."  

Finally, other normative considerations creep into carcinogen risk assessment because of the uncertainties which currently pervade it, but I will not develop those in great detail here since they are somewhat more peripheral to this discussion.  

6 Carl F. Cranor, Regulating Toxic Substances: A Philosophy of Science and the Law 25-28 (1993) (For example, risk assessments are substantially influenced by normative judgments in choices of models, and the cumulative efforts of many typical
What can we infer from the normative presuppositions in risk assessment? First, because of pervasive uncertainties in risk assessment, it differs from core areas of science. There are many more uncertainties in risk assessment than in more settled scientific areas. Of course, cutting-edge science has many uncertainties, but because of the contentious nature of many risk assessment issues, the significance of the uncertainties will probably be magnified. Second, because of the response to uncertainties and other considerations noted above, risk assessment is substantially permeated with normative judgments. Third, as developed in some detail elsewhere, it is difficult or impossible to separate risk assessment from risk management. For one thing, as noted above, where statistical studies are needed to provide evidence of toxicity, decisions about sample size, cost of experiment, and the desired degree of accuracy all involve normative decisions in the very conception and design of such studies. Moreover, since risk assessment is an imperfect procedure, there will be mistakes and which mistakes the process is designed to avoid is an important normative issue. One must make a policy decision (or decisions) about the degree of accuracy sought, the social costs of achieving that accuracy and the importance of the risks to be prevented. Explicit discussion of alternatives to conventional risk assessment and risk management practices of scientists should occur because many current practices may frustrate preventive goals.

Fourth, as just noted, there is often a logical incompatibility between low false positives and low false negatives; in other cases there are scientific practices that may frustrate the discovery of risks. In such circumstances, scientists need to be wary of their rigorous commitment to low false positives or to certain practices precisely because these can inadvertently prevent the discovery of the risks in question.

Fifth, a larger point is that evidentiary values implicit in scientific inquiry (low false positives, a scientific notion of accuracy, a particular conception of rigor and a desire not to add mistakenly to the stock of scientific practices may greatly slow risk assessments).

scientific knowledge) may inadvertently trump the values the risk assessment serves. The enumerated scientific practices may inadvertently elevate the virtues of a scientific procedure above the normative concerns that motivate assessments, can inadvertently influence outcomes and possibly prevent or delay discovery of the very harm for which the assessment was undertaken. Very persuasive reasons would have to be provided for research scientific evidentiary goals always to trump such social goals. For example, an insistence on very low false positive rates may prevent the discovery of existing risks for which the study was done. By contrast, an approach to risk assessment more sensitive to the context would avoid insensitive reliance on low false positive rates, practices which would frustrate the discovery of the risks in question, a misplaced sense of rigor, and a misguided notion of accuracy. In some circumstances where statistical studies are the primary evidence this might mean tolerating somewhat higher false positives rates so that false negative rates could be lower.

Sixth, if the aim of risk assessments should be to reduce the total social costs of mistakes together with the social costs of the procedures for doing the risk assessment, there is no guarantee that scrupulous scientific accuracy will minimize the social costs of the procedure. An abstract point is that scientific accuracy is only one of several variables that would help minimize mistakes and there is no assurance that vigorously pursuing it will result in the lowest total cost of mistakes and the procedure as a whole. A secondary reason is that the scientific paradigm tends to focus on avoiding false positives and on having a great deal of confidence in one’s positive results. But as already noted this commitment may actually frustrate some of the public health goals of the risk assessment process itself. For example, economists have noted that one of the more scientifically rigorous procedures for identifying carcinogens, the use of the animal bioassay, is rarely worth the cost for the evidentiary information it provides.\textsuperscript{8} Depending upon the social costs at stake, a less rigorous adherence to scientific evidentiary standards may better reduce the total social costs of mistakes.


\textsuperscript{8} \textit{Risk: Health, Safety & Environment} 123 [Spring 1997]
Recognizing that the aim of risk assessment should be to reduce the total social costs of two kinds of mistakes plus the costs of the procedures releases us from a particular rigidly scientific paradigm. Freeing us from that then opens up opportunities to develop different procedures and models for assessing the risk from toxic substances. (I return to this below.) Further, the costs of the procedures themselves can be substantial. In fact, I believe a scientific-intensive paradigm which focuses on the case-by-case evaluation of the toxic properties of each substance with the further aim of understanding the mechanisms by which the toxic properties operate (a quite appropriate procedure for doing good science and good toxicology) can be costly in terms of money, time and human resources. Thus, the aim of reducing the total social costs of doing risk assessments invites us to reevaluate existing procedures. If we can achieve much the same aims with an overall reduction of money, time and other costs, this revised picture of risk assessment would recommend that we adopt the less socially costly procedures. This suggests that one needs the idea of the minimum kind and amount of evidence to judge the toxicity of substances to serve relevant social goals. Thus, for example, for warning people that substances might be toxic, so that they can take protective steps, one might need a less substantial evidentiary base than for setting ambient concentrations of a substance that would reduce risks to some socially acceptable level. The evidentiary basis needed for technology forcing statutes might lie between these extremes.

There can be social costs from slow risk assessments when regulators are working under either post-market or premarket regulatory statutes. Post-market statutes leave people exposed to toxic substances in commerce until their toxicity is discovered and they are regulated. The costs from slow procedures are the monetary and social costs of the procedures themselves and opportunity costs resulting from the slow rate of assessment which may leave unevaluated other toxic substances in commerce to inflict injuries. Inaction is costly, thus, the rate of the identification and assessment of carcinogens is an important normative consideration. Similarly, even under premarket statutes there can be costs to the rate of evaluation. Under a premarket statute, a substance proposed for introduction into commerce is not harming anyone, if it
Cranor: The Normative Nature of Risk Assessment

131

turns out to be toxic, until there is actual exposure, so the slowness of a premarket procedure does not pose the same problems as in the postmarket case. However, if, in an extreme case, a substance will bring lifesaving benefits, slow evaluation of its toxicity may result in serious health consequences from inaction. For example, if certain AIDS drugs would in fact save lives; slow premarket procedures might substantially harm AIDS victims. In such cases, it seems appropriate to have expedited reviews and possibly a different evidentiary base. However, there is a cautionary note. In the past, different parties have complained that both postmarket and premarket procedures have been too slow. Sometimes parties to the debates have argued in effect for keeping postmarket procedures science-intensive, thus slow, while expediting premarket procedures. However, these have tended to be self-interested arguments. A more principled view would be to design both sets of procedures so that they better serve the aims of reducing the total social costs of identifying and assessing toxic substances and protecting the public health.

To illustrate the consequences of reducing total social costs, consider alternative procedures to existing carcinogen potency and identification procedures. Current carcinogen potency assessments are typically slow, taking from 0.5–5 person years per substance. While these are as accurate as they can be at present (insofar as this can be determined), they are expensive, have substantial opportunity costs, and do not address well public health costs posed by the whole universe of known carcinogenic substances or the whole universe of chemical substances. There are 72,000 substances in EPA’s inventory of chemicals under the Toxic Substances Control Act, most of which are

9 Information presented to California’s Proposition 65 Science Advisory panel on which the author served.
10 Science-intensive carcinogen potency assessments for 200 substances would cost from $7–70M whereas one expedited procedure evaluated in California costs about $4,000 and another $40,000 (Information presented to California’s Proposition 65 Science Advisory Panel). Also, science-intensive assessments require the expertise of one or more highly-trained toxicologists, whereas the two expedited procedures necessitate much less reliance on expensive experts.
11 California Environmental Protection Agency, Procedure for Prioritizing Candidate Chemicals for Consideration Under Proposition 65 by the “State’s Qualified Experts” (over 500 identified by the State of California).
12 U. S. Congress, General Accounting Office, Toxic Substances Control Act:

8 Risk: Health, Safety & Environment 123 [Spring 1997]
But risk assessments can be done differently. We could use faster procedures that might be slightly less “accurate” but would decrease the number of known but unassessed carcinogens. This would result in lower overall social and agency costs. In fact there are technical assessment procedures, scientific approximations really, which are quite accurate and relatively quick that can be used for many regulatory purposes. These results are described elsewhere and have been adopted into regulation in California. Such procedures permitted the potency assessment of 200 substances in less than one year by the California Environmental Protection Agency and added more than 140 substances to the list under California’s Proposition 65. These are not the only expedited procedures that have been considered for public health protection. In 1989, the Occupational Safety and Health Administration proposed to use American Council of Industrial Hygienists standards to set exposure standards for 428 air toxics to greatly expedite its regulatory agenda and other researchers have proposed somewhat different expedited regulatory procedures.

Similar procedures could perhaps be adopted for the identification of carcinogens. Currently, carcinogens are identified for regulation either by means of human epidemiological studies or animal bioassays. Both are costly and time-consuming, but given human health concerns, animal bioassays are the preferred procedure, since it seems wrong to wait for a sufficient number of human deaths to be detected by epidemiological studies before taking precautionary action. However, even animal studies are slow, taking from five to seven years to complete, and costly (about $2M per study). While such studies are being conducted and analyzed other substances are being ignored — there are not enough toxicologists to assess all suspect substances in this


13 The Environmental Protection Agency has evaluated only 2% of 62,000 chemicals in commerce when the agency began to assess new substances under the Toxic Substances Control Act. Id. at 3.


15 See Air Contaminants Standard Overturned, Occupational Health and Safety Newsletter, July 8, 1992, at 103.

kind of detail. Thus, like potency assessments, at present the identification of carcinogens is slow, is resource-intensive and appears to have substantial opportunity costs.

The paradigm shift suggested above, thus opens opportunities for identification procedures. If there are expedited methods that keep false negatives quite low (the more important concern from a public health perspective it appears), keep false positives within a reasonable range and lower identification costs, we should prefer them to current procedures because of lower total social costs and how well they serve overall public health and other values. Some such procedures have been identified, and other preliminary work suggests the following.

There are various short-term tests (STTs), e.g., of mutagenicity, structure-activity relationships, etc., that are quick, inexpensive, and relatively easy to use compared with animal carcinogenic bioassays. Their use could provide an approximate characterization of the risks posed by substances and result in public health benefits (because toxic substances would be identified earlier). A recent workshop evaluated the accuracy of approximately fifteen different expedited procedures for their accuracy against the results of animal bioassays. Validating them faced some difficulties, but nevertheless several procedures appear to be sufficiently accurate to use for identification purposes. In addition, modeling procedures similar to those used by environmental economists, suggests that even STTs that are less than fully accurate, with false negative and false positive rates above .05, may have uses both in premarket and in postmarket regulation, depending upon some facts about the world and what is at stake. For example, if the percentage of carcinogens in the chemical universe is 7.5% or greater, then STTs appear to be justified compared with reliance upon the results of animal bioassays for providing preliminary identification of carcinogens. Such

17 Lester Lave & Gilbert S. Omenn, Cost-Effectiveness of Short-Term Tests for Carcinogenicity, 324 Nature 29 (1986).
procedures will result in mistakes, both false positives and false negatives, but the balance of the social costs of such mistakes favors using faster, even if somewhat less accurate identification procedures, rather than slow science-intensive methods. The greater the percentage of carcinogens, the better the case for utilizing STTs.²⁰

Lists of substances identified as carcinogens by expedited procedures might have uses both outside and inside regulatory processes. Outside the regulatory environment lists of likely carcinogens based upon STTs could be used by academics in their research to begin to fill data gaps, provide clues to the toxicity of substances for further investigation, or help confirm or refute preliminary toxicity implications. Such tests and lists could also be used by private industry for its own internal considerations (as they almost surely are) or by the public for self-protection, a kind of informal analog to California’s Proposition 65. In a regulatory context, in addition to the above uses, such lists could be used by governmental agencies to trigger testing requirements under the Toxic Substances Control Act, used in conjunction with listing statutes such as California’s Proposition 65, Superfund Toxic Release Inventories, or other such lists which have some legal consequences, but which do not require extensive expenditures for money for reducing exposures to the public. They could also be used to provide information for pollution-reduction, and perhaps used in conjunction with technology-forcing statutes to identify likely carcinogens and then to require industry to use technology to reduce exposures to them. There might even be some uses in very limited circumstances with statutes requiring the regulation of ambient exposure levels of particular substances although such uses of expedited procedures would be the most controversial.

Each different use would likely require different levels of evidentiary support, but that is point. Conceiving of risk assessment as partly normative, a policy tool, opens up the possibility of different evidentiary approaches for different informal and regulatory purposes — just as we otherwise demand different kinds and amounts of evidence in different social and legal contexts.

Eighth, the presence of normative considerations are an asset, not a liability, in risk assessment. Recognizing the normative aspects not only frees us from certain conceptions of the field, but it also can be used to aid in decision making and in reconceiving the procedures. Elsewhere, I have suggested how normative considerations can be used as they are in the law to predispose risk assessments to decision outcomes based on the best readily accessible scientific information and policy goals in the face of pervasive uncertainty and in the absence of better information.\textsuperscript{21} If we conceive of risk assessment procedures somewhat less like a strictly scientific activity and somewhat more like those of other institutions, such as the law, normative considerations are both more appropriate and have greater pride of place, instead of being treated like crazy, unwanted distant relatives.

Finally, the above discussion suggests a further point about the virtues of those charged with identifying, assessing and regulating toxic substances. Typically risk assessors are trained in scientific programs, many in toxicology, and are apt to recognize the virtues of research science: the importance of understanding phenomena, the need to do proper toxicology studies and even the need to understand underlying mechanisms by which a substance works. This is as it should be for research. However, when one enters the risk assessment and risk management arenas, additional virtues are appropriate.

Risk assessors should acquire some of the virtues of public health officials (for that is in part what they are). A public health official also has scientific training but may have to act on the best available information without having the luxury of studying an issue until he or she understands it with a high degree of certainty (there can be costs from the careful evaluation of the substance in question, if people remain exposed, and other substances are ignored). A public health official will approach the problem of small pox much differently than will a Louis Pasteur — one must act to protect the public, the other seeks to understand phenomena. Both are needed for infectious diseases and for addressing the problems posed by toxic substances, but their roles should not be confused.

Risk assessors should also acquire some of the virtues of engineers. These are appropriate because engineers must frequently solve problems even though they do not have full understanding of phenomena in question. For example, airplane propellers kept airplanes aloft before their aerodynamic properties were fully understood.22 The point? Engineers frequently must act on the basis of clues about how a problem should be solved and must rely upon scientific approximations and procedures to accomplish their aims. They also may address unknowns or uncertainties by means of approximations and safety factors, since they cannot wait for full theoretical scientific understanding to guide their work.23 Thus, the virtues of research scientists who become risk assessors and risk managers might appropriately be augmented by some of the virtues of both public health officials and engineers. Risk assessors, thus, would be disposed to use readily available scientific evidence, assess it in a timely manner, and use appropriate scientific approximations to protect the public health. A person with the combination of research, public health, and engineering virtues might be in a better position to appreciate the need to minimize the total social costs of these imperfect procedures and to implement them.

23 Id. at 167.