1995

Direct Inference, Probability, and a Conceptual Gulf in Risk Communication

Vern R. Walker
Maurice A. Deane School of Law at Hofstra University

Follow this and additional works at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship

Recommended Citation
Available at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship/512

This Article is brought to you for free and open access by Scholarly Commons at Hofstra Law. It has been accepted for inclusion in Hofstra Law Faculty Scholarship by an authorized administrator of Scholarly Commons at Hofstra Law. For more information, please contact lawcls@hofstra.edu.
Direct Inference, Probability, and a Conceptual Gulf in Risk Communication

Vern R. Walker 1

Differences in the conceptual frameworks of scientists and nonscientists may create barriers to risk communication. This article examines two such conceptual problems. First, the logic of "direct inference" from group statistics to probabilities about specific individuals suggests that individuals might be acting rationally in refusing to apply to themselves the conclusions of regulatory risk assessments. Second, while regulators and risk assessment scientists often use an "objectivist" or "relative frequency" interpretation of probability statements, members of the public are more likely to adopt a "subjectivist" or "degree of confidence" interpretation when estimating their personal risks, and either misunderstand or significantly discount the relevance of risk assessment conclusions. If these analyses of inference and probability are correct, there may be a conceptual gulf at the center of risk communication that cannot be bridged by additional data about the magnitude of group risk. Suggestions are made for empirical studies that might help regulators deal with this conceptual gulf.

KEY WORDS: Direct inference; probability; risk communication; risk perception; uncertainty; variability; personal risk.

1. INTRODUCTION

Risk communication sometimes takes place across a conceptual gulf between two groups with different perspectives. On the one hand, regulators and public health scientists make predictions about the incidence of injury or disease within groups. On the other hand, members of the public are often concerned with the probability of injury for specific individuals. This article explores certain confusions caused by the differences between these conceptual frameworks. The discussion is divided into two major issues: (1) the direct inference from statistics about groups to probabilities for specific individuals, and (2) different interpretations given to such probability statements. The main thesis is that the conceptual gulf is fundamental and has a logical basis, and might not be bridged, or even significantly narrowed, merely by additional empirical studies about group risks. This conceptual difference has important implications when designing empirical studies into risk perception and risk communication, and suggests several hypotheses for future research.

2. DIRECT INFERENCES FROM GROUPS TO INDIVIDUALS

2.1. Groups vs. Individuals

Regulators and the public health community are naturally concerned with characteristics of groups. First, they typically rely upon data from groups in arriving at a risk assessment—whether the data are from animal studies or epidemiologic studies. Second, they use such data to make predictions about groups, by using either statistical models (estimating population parameters) or...
probability models (combining statistical data with causal theories). Third, action alternatives are often evaluated by their effects upon different groups of potentially affected persons. Finally, the regulatory agency naturally conceives of its political constituencies as organized into groups having a commonality of interest.

Due in part to this natural orientation toward groups, regulators and public health scientists tend to avoid making inferences from information about groups to conclusions about specific individuals. By contrast, consumers of risk information do attempt to reach conclusions about specific individuals based on statistical information about groups. When I acquire risk information about groups (e.g., Americans who do not smoke, males over age 40), I try to use it to reach conclusions about risk to myself or to specific others (e.g., my father). Philosophers have called this “direct inference,” and have studied the logical problem of determining when there is sufficient information to warrant the direct inference. For example, if 5% of people in a group are likely to incur a particular injury, what more do we need to know before we are warranted in concluding that a specific individual in that group has a 0.05 probability of suffering that injury?

2.2. A Data Gap in Individuating Risk Factors

Several philosophers, representing different theoretical camps, have recently argued that a prerequisite for a direct inference from group to individual is either complete knowledge or “total evidence.” The key requirement is what I will call “complete identification.” What is required is that the only identifying information about the individual that we will take into account is the fact that he or she is a member of the relevant group. That is, the inference from group statistics to individual case is warranted only if we completely identify the individual as a member of the group. In the above example, the inference from a statistic about 5% of the group to a 0.05 probability about a specific individual is warranted only if no additional information about the individual is considered relevant, other than group membership. This “complete identification” requirement can be met by making the inference apply to a “statistical individual,” a hypothetical member of a statistical population from which a sample is drawn.

We reach conclusions about statistical individuals whenever we refer to the “median individual” or the “99th-percentile individual.”

A second logical requirement for such a valid direct inference is “equiprobability”—a conceptual cousin of the classical “principle of indifference.” We can transform the statistic about the group (“incidence = 5%”) into a probability about a specific person (“probability = 0.05”) if we regard the selection of the subject individual as a “simple random draw” from the group. If every individual in the group has an equal chance of being selected, then any one selection will have a 0.05 probability of drawing an individual who will suffer injury.

Thus, we justify the direct inference, first, by completely identifying the individual as simply a member of the group and, second, by assuming equiprobability of selection among group members.

These two logical requirements, however, are really simplifying assumptions. A complete logic of individuation would begin with an estimate of a relevant group risk (e.g., a statistic such as incidence) and adjust that estimate up or down depending upon individuating risk factors, thus arriving at a total risk to the specific individual (expressed by a probability). For example, if the national statistics for automobile fatalities were to yield a lifetime risk of 1 per 100 persons of dying in an automobile accident, I should not automatically infer that my probability of dying in a car accident is 0.01. I would need to take into account such individuating factors as my driving skills, my geographic area of the country, the time of day that I drive, and my pattern and extent of automobile use. Other examples of individuating risk factors are the genetic, developmental and environmental factors that make my body react in a particular way to a particular environmental toxin. Biological variation is the manifestation of such individuating causal factors.

Without such individuating information to tailor the generic risk to the specific case, the direct inference is based on an evidentiary data gap that is often filled by simplifying assumptions about “complete identification” and “equiprobability.” By reaching conclusions only about statistical individuals and by assuming that every specific individual has an equal chance of being those statistical individuals, regulators avoid the need for individuating risk information, as well as the difficult problem of determining the direction and magnitude of the individuating adjustment. A specific individual might have scores on independent variables that lie outside the range actually studied, might possess characteristics that are probably relevant but on which there are no data, and might have other characteristics whose nature and relevance are completely unknown. Moreover, there might be interactive effects between relevant factors that will not be studied in the foreseeable future. Peculiarities of genetic makeup, developmental process, and environmental history combine to make simplifying assumptions desirable.
In theory, we could investigate the individuating risk factors. We could design studies around an individual's particular characteristics, using subjects whose genetic compositions, and developmental and environmental histories, are similar to that individual's. Although we would still have to qualify our conclusions to reflect inherent methodological uncertainties, we could in theory reach conclusions about risk that would be much more tailored to the particular individual. But it is clear why society cannot and will not undertake such a research program. There are problems with technological, ethical, and economic feasibility, as well as with political acceptability. Witness the difficulties in obtaining funding to study even distinct subpopulations, let alone specific individuals. The public health community must set its research priorities based on such considerations as the importance to the society of the type of injury or illness, the prevalence in the population, the susceptibility of risk factors to control, the cost of generating information, and other competing demands on societal resources. The only acceptable social solution is to explain and predict individual cases only indirectly, by assuming that they are modeled adequately by feasible studies about groups. But this reality sets up an inherent tension within risk communication.

2.3. Resulting Tension in Risk Communication

In risk communication, the societal group perspective and the individual perspective inevitably meet. On the one hand, regulators and scientists are trying to convey to individuals the group information that is available, to use that group information to make decisions affecting public health and the environment, and to obtain the consent and acceptance of individuals for societal projects. The regulator might also assume that the group information supports inferences to individual cases, and that potentially affected individuals should simply translate group risks into probabilities for themselves. Regulators sometimes imply that individuals who refuse to do so are acting irrationally.

Individuals, on the other hand, are to some extent trying to use this group information to reach warranted conclusions about the risks to themselves as specific individuals, to their family members, and to their neighbors. Although scientists attempt to estimate variability in the face of uncertainty, the individual is trying to determine where a specific individual is in the variability distribution. When a regulator makes a best estimate of expected incidence of “one in a million,” a member of the public may be wondering about the probability that a specific individual is that one person in a million.

In addition, a potentially affected individual knows that he or she has not been specifically studied. He or she intuits the existence of a data gap concerning individuating risk factors. Epidemiologic studies do not study the specific individual, and experts debate whether animal results are predictive even for groups of humans, let alone for specific individuals. The rational individual may well conclude, depending on the nature of the hazard, that no one knows how well a particular case has been modeled by the group data or how any specific individual will be affected.

Furthermore, individuals may resent the regulatory simplifying assumptions discussed above. They might resent treating individuals as fungible, as mere members of a group, and resent assuming that their specific chances of being adversely affected are the same as those for everyone else. They want to know what the chances actually are for specific individuals.

Specific individuals may therefore take seriously the lack of knowledge concerning individuating risk factors. The fact that the inference gap is logical in nature suggests that it is not irrational to refuse to make the simplifying assumptions so essential to regulators. At most, I as an individual might be charged with undue caution in my decision about how to “deal with” that gap. But I may be acting very rationally in refusing to assume it away.

The specific individual, therefore, will probably employ personal decision rules for how to proceed in the face of ignorance, just as regulators develop policies for proceeding under uncertainty. Several points can be made about such personal decision rules. First, although the total individual risk might be lower than the group risk, it is probably more common that the direction and magnitude of adjustment due to the individuating risk factors are unknown. This would be the case with those environmental toxins for which we have incidence predictions, but know very little about the causal mechanism and the role of individuating risk factors. In cases involving such ignorance, it might be reasonable for an individual to act as though her total individual risk is higher than the regulator's conservative estimate of group risk.

Second, a reasonable decision about how to deal with a data gap for individuating risk factors probably takes into account the expected distribution of benefits from any societal decision. If the individual is certain to benefit personally by the societal decision, then it might be reasonable to tolerate greater uncertainty about the individuating risk factors.
Third, a reasonable decision rule might postulate a right of the potentially affected individual to apply his or her own personal decision rule. If there is a need to deal with a data gap concerning individuating risk factors, or a need to balance possible benefits against such risks, then obtaining the informed consent of all potentially affected individuals might play a major role in the personal decision rules of many people. I may be willing to make various decisions affecting myself without understanding my individuating risk factors, but I might not be willing to impose that personal decision rule on others.

2.4. Implications for Empirical Studies

The existence of a data gap with respect to individuating risk factors, and its logical nature, may cause inherent tension within risk communication between regulators and the public. This difference in perspective has implications for empirical studies concerning risk. Instruments to measure risk perception should be sensitive to the difference between an individual's estimate of the group risk and his or her estimates of risk to specific individuals. A number of studies have investigated differences between experts and the public in the perception of risk, and some have addressed differences in risk perception as a function of the "target" of the risk (self, family, people in general). (10)

Individuals might use different decision rules to deal with data gap about individualizing risk factors, depending upon whether it is a "first-person" risk (to themselves, or to close family members or others with whom they identify) or a "third-person" risk (to strangers). (10, 11) This difference in decision rules is not the same phenomenon as "not-in-my-back-yard": a person in Maine might be opposed to a proposed federal nuclear waste repository due to his perception of individualized risk to an individual in Nevada. We should design survey instruments keeping in mind such differences in perceived individualized risks.

In addition to risk perception studies, empirical studies of risk communication should also recognize the distinction between group risk and individual risk. For example, researchers increasingly study "trust" as an explanatory variable for human acceptance of regulatory risk estimates. (12, 13) But the problem of estimating total individual risk in the face of a data gap about individualizing risk factors is a matter of logic that cannot be eliminated by trust. Even honest officials may be unable to fill the individualized data gap with adequate information. If the problem for the individual is that no one can predict what will happen to him, then this is not a communication problem that can be resolved simply by increasing good faith or trust.

Of course, the failure of regulators to recognize and acknowledge the logical problem could increase public distrust. Part of public distrust might be due to frustration that regulatory officials do not acknowledge their ignorance about the individualized risk factors at work in the specific case. The public is left to wonder whether officials do not see the logical problem at all, or whether they are wilfully ignoring it. It is an empirical question how techniques of acknowledging ignorance might influence public distrust.

3. TWO INTERPRETATIONS OF PROBABILITY STATEMENTS

3.1. Objectivist vs. Subjectivist Interpretations

The communication difficulties caused by the inference gap between group risk and individualized risk are often compounded by a difference in the meaning of probability statements. Experts and non-experts can interpret probability statements very differently, and this can also contribute to confusion within risk communication.

It is traditional, and adequate for purposes of the present discussion, to divide interpretations of probability statements into two broad approaches or schools: objectivist interpretations and subjectivist interpretations. (4, 14-16) The disagreement is over what a probability measures. A standard objectivist interpretation is that a probability statement is about a relative frequency of occurrence in the long run. (17) For example, the probability of dying in an automobile accident estimates the fatality rate in the relevant population. The objectivist interpretation emphasizes the frequency of the events in which we are interested, and the nature of the causal systems producing those events.

By contrast, a typical subjectivist interpretation is that a probability indicates the degree of confidence that someone has or should have that an event will occur. (4, 15, 16) Probabilities are about the "subjective" mental states of actual people or hypothetical reasonable people. The probability of dying in an automobile accident might be thought of as expressing the "betting odds" that such a fatality will occur. The subjectivist interpretation emphasizes not the frequency of events, but the limitations of our knowledge about those events,
and our resulting lack of confidence in making predictions.

3.2. Probabilities for Unique Events

Objectivist and subjectivist interpretations of probability are usually compatible, and they can coexist in an ambivalent tolerance. An exception to this complementarity, however, is the prediction of a single, unique event. Objectivist theories have always had difficulty interpreting probability statements about unique events occurring to specific individuals. What objectivist meaning can there be to the assertion that I personally have a 0.5 probability of developing a certain kind of cancer within 5 years? My getting cancer or not within the next 5 years is not a repeatable event, so “relative frequency of occurrence” has no obvious meaning. An objectivist is tempted to resolve this problem by interpreting the probability as “really” about cancer incidence in groups of “similar people.” Thus, a probability statement about my developing cancer is really a statement about the expected incidence of cancer in a group with relevant characteristics similar to mine.

A subjectivist, by contrast, has no difficulty at all interpreting probability assertions about unique events, since all probabilities express degrees of confidence that an event will occur. A probability of 0.5 expresses the mental state of equipoise, in which the predictor has no more confidence in the prediction than in its negation. From the standpoint of confidence, it matters not at all whether the prediction is about a specific individual or about a group.

3.3. Resulting Misunderstandings in Risk Communication

Many regulators and public health scientists naturally intend an objectivist interpretation for their probability assertions. First, since regulators usually rely on empirical data about group incidence and severity, they also tend to think of their conclusions in terms of frequency. For example, from tumor incidence in a sample of test animals they might infer a worst-case incidence in a human population. Second, they tend to believe that disease occurs with cause-and-effect regularity, and that most causal processes produce their effects regardless of human degrees of confidence. Moreover, in communicating risk to the public, they might see themselves as trying to replace uninformed, subjective fears with objective information. For all of these professional reasons, they tend to interpret probabilities objectively.

But the member of the public who wishes to know what is going to happen to a specific individual is concerned about unique events. “Will this nuclear reactor ever melt down?” “Will my daughter develop cancer due to exposure to this chemical?” This nuclear reactor will either melt down or not, and my daughter can live only one future. I accept probability assertions about the future largely because I understand the limitations of human knowledge, and because I am intuitively aware of the data gap at the heart of the direct inference problem. But these “facts of life” also encourage me to interpret probability statements in a subjectivist fashion, as expressing the speaker’s degree of confidence in the prediction.

In risk communication, the objective and subjective interpretations of probability statements can clash in confusing ways. For example, an objectivist-oriented regulator might interpret a “one in a million” risk as a worst-case estimate of expected additional cancer incidence in the exposed group. Statistical devices that characterize the uncertainty around that prediction, such as confidence intervals, do not change the nature of the prediction: it is still an estimate of frequency of occurrence. The objectivist communicator understands the probability as predicting a very small incidence (worst case), and expects the public to be reassured by this fact.

A member of the public, however, might interpret the “one in a million” probability in a subjectivist way, as indicating the regulator’s high degree of confidence that no one will develop cancer. But if that is what the probability means, the individual may be puzzled by the basis for such high confidence. First, even such high confidence in group predictions is questionable, especially if the evidentiary basis consists of animal data showing that the chemical agent can sometimes cause cancer in animals. Second, there may appear to be no basis for confidence in a prediction for any specific individual, given the problem of direct inference. Even if the best group prediction is in fact “one in a million,” the absence of individuating data should lower any subjective confidence that I will not be the affected individual. If the probability assertion expresses degree of confidence, not relative frequency, then it may be difficult to understand the basis for the regulator’s apparently high degree of confidence.

If there can be sound logical reasons for an individual to resist making the direct inference from group risk to individual risk, this inference problem can be compounded by differences in interpreting probability statements about risk. By making a probability estimate
of "one in a million," a regulator may intend to make the well-founded (objectivist) claim: "I can't be very confident at all, but my best prediction based on limited evidence is that at most one or two additional cancer cases will occur per million people in the population."

But a member of the public hearing this probability may interpret the low probability estimate as expressing a high degree of confidence on the part of the regulator that no individual will develop cancer from the exposure. The public might hear the unfounded (subjectivist) statement: "I'm very, very confident that you won't get cancer from this exposure." These are very different interpretations, and can generate a great deal of risk communication. As a result, the regulator might perceive the public as irrational and distrusting, and the public might view the regulator as acting from bad faith, naivety, or ignorance.

4. ADDRESSING THE CONCEPTUAL GULF IN RISK COMMUNICATION

The two conceptual problems discussed in this article are separate and distinct. The inferential gap from statistics about groups to conclusions about individuals is an evidentiary problem, while the objectivist-subjectivist distinction concerns the meaning of probability statements. The evidentiary gap concerning individuating risk factors can be viewed in either objectivist or subjectivist terms, and objectivists and subjectivists must both address the direct inference problem. The compound difficulty for risk communication is that both of these conceptual problems intersect when a member of the public is worried about probabilistic assessments of the risk to specific individuals. This combination creates a complicated conceptual gulf between the parties to the communication. The lack of information about individuating factors combines with the confusion about the meaning of probability statements to produce miscommunication. Moreover, this source of miscommunication complicates such additional problems as misunderstanding scientific information, misperceiving group risk, and mistrusting public officials.

Despite the confusion, it is important to attempt to separate this conceptual gulf from the other types of problems. This gulf might not be bridged, or even significantly narrowed, by additional data or empirical studies about the magnitude of group risk. Additional studies about causal risk factors usually change only the group statistics, without significantly narrowing the direct inference data gap. Lowering the point estimate of future group incidence, or contracting the confidence interval around that estimate, might not reassure the at-risk individual.

It is possible that research expressly addressing this conceptual gulf might lead to more successful strategies for dealing with the problem. First, although some studies have touched on the extent to which the direct inference gap actually influences a population's thinking about risk, this effect should be studied explicitly. The suggestions for study design made in Section 2.4 above would be relevant for such studies. Most members of the public are concerned to some extent with both the total risk to the community and the individualized risk to themselves or specific others. Empirical studies measuring the relative importance of these two perspectives in public thinking might show that questions about individualized risk sometimes play a psychologically minor role. I suspect, however, that we will find that worries about individuation predominate among individuals who are personally at risk. Such worries might also predominate when biological variability plays a greater role in producing an effect (for example, susceptibility to a weak environmental toxin emitted from a waste incinerator) than when a single set of physical forces drives outcomes (for example, in a nuclear reactor meltdown).

Second, it might be possible to develop analytical techniques to extract more individuating information from group studies. Perhaps summary statistics about groups can be supplemented with integrative techniques centered on individuals. While research into distinct subgroups—such as women, children, and racial subgroups—are movements toward the individual, we have a long way to go down that road. Although we will never completely close the logical gap between groups and individuals, we may be able to use available data to provide risk information to the public in a way that assists them to make direct inferences to specific individuals.

Third, studies might be undertaken to investigate the types of personal decision rules that people actually use to deal with data gaps concerning individuating risk factors. This information might lead to suggestions for providing information to the public that they will regard as useful and responsive to their decisional needs.

Finally, it should be useful to study more effective means of communicating probabilistic information, taking into account the different ways of interpreting probability statements. This problem is, after all, a problem of meaning, rather than a gap in evidence. The task is learning how to communicate relative frequency estimates, while at the same time communicating the in-
tended degree of confidence in those estimates, without getting these two tasks hopelessly confused.

Any adequate solution for risk communication will have to recognize that the conceptual differences discussed here are real and fundamental. Individuals are often acting very rationally in refusing to apply directly to themselves the regulator's risk estimates for a population. And if the public is being rational, then trying to "educate" them out of being so is not only doomed to failure, but probably unethical.

REFERENCES