RISK: Health, Safety & Environment (1990-2002)

Volume 9 | Number 1 Article 5

January 1998

Risk Regulation and the Faces of Uncertainty

Vern R. Walker

Risk Regulation and the "Faces" of Uncertainty

Vern R. Walker*

Introduction

The problem of "decision making in the face of uncertainty" is how to regulate on the basis of incomplete information that has significant potential to be inaccurate. My primary goal is to clarify some aspects of such decisionmaking through surveying types of uncertainty inherent in information. I present a taxonomy for kinds of inherent uncertainty — a classification and description of the "faces"of uncertainty. A second goal is to identify which kinds of uncertainty can and cannot now be measured quantitatively.

A contextual assumption is that we need to evaluate decision rules for dealing with faces of uncertainty. As a social enterprise, risk regulation, whatever its substantive objectives, should be as *effective, efficient* and *equitable* as practically possible. These "three E's" form a set of "process objectives" or "meta-goals."

First, risk regulation should be as effective as possible to maximize the likelihood of achieving a selected balance of risk taking and avoidance. Second, with regard to efficiency, we may, e.g., try to minimize adverse collateral consequences, maximize net benefits after accounting for costs, or use available resources cost-effectively. Third, the goal of equitable results is partly a concern for affected parties' appropriate distribution of costs and benefits, and partly a respect for the potentially affected parties' substantive and procedural rights. While these three meta-goals are distinct and complementary, they are also sometimes in tension. The optimal and most achievable blend of all three meta-goals depends upon circumstances surrounding each regulatory action.

These meta-goals form an evaluative context for risk regulation and derivatively for regulatory factfinding. A necessary condition for

* Professor of Law, Hofstra University School of Law. He holds a B.A. from the University of Detroit, an M.A. and Ph.D. (Philosophy) from Notre Dame University and a J.D. from Yale Law School. Email: *lawvrw@hofstra.edu*.

achieving the optimal blend of effectiveness, efficiency and equity for a particular regulatory action is the appropriate use of available scientific information. For risk regulators to predict consequences accurately, to minimize collateral side-effects, and to allocate fairly both costs and benefits, scientific information and conclusions supporting a regulation must be adequate and accurate. We can achieve appropriate accuracy by understanding the types and degrees of uncertainty inherent in scientific information.

This idea is illustrated by considering the role three principal categories of causal information needed to justify and guide regulatory decision making: information about risks, benefits and costs. Information about risks associated with a regulatory action is essential. Such information is often the result of quantitative assessment, grounded in information about causal processes. Information about possible benefits is perhaps less studied as a separate category, except by economists, but we can imagine a methodology for "benefit assessment" with or without attempts at economic valuation. Finally, cost analysis is also based on causal information about both natural and market processes. The point here is not to defend a comprehensive classification scheme for substantive types of information, but to illustrate that in each category, information about causation is of fundamental importance. We need to know *what causes what* to determine risks, benefits or costs — both their magnitudes and their likelihoods of occurrence. If our information about causal action has inherent uncertainties, then risk, benefit and cost assessments are subject to those uncertainties, and any justification of the regulation as efficient and equitable is subject to uncertainty.

The evaluative context of risk regulation sets the stage for the article's primary task, describing a taxonomy of the kinds of uncertainty inherent in causal information, and therefore indirectly inherent in conclusions about risks, benefits and costs. From the uncertainty standpoint, causal information is divided usefully into two major categories: information about groups and information about individuals. An example of causal conclusions about groups is that exposure to a certain chemical through inhalation can cause human cancer and that the upper-bound risk of excess cancer cases in groups

exposed at certain levels is estimated as one-in-one-million. Examples of causal conclusions about individuals are that my daughter probably will not develop cancer from her exposure to the aforesaid chemical or that a specific individual's case of lung cancer was probably from cigarette smoking. Each category, group and individual, has its own distinctive types of inherent uncertainty. These are logically distinct, generally independent of each other, and cumulative in contributing to the "total uncertainty" inherent in ultimate conclusions about causation. I will build up a profile of these types of uncertainty in a series of steps, beginning with the uncertainties inherent in group conclusions.

Information about Groups

There are at least five distinct types of uncertainty inherent in causal information about groups: uncertainties due to concept selection, measurement, sampling, mathematical modeling and causal modeling. I will discuss each of these in turn.[1]

Concept Uncertainty. The first kind of uncertainty inherent in information about groups stems from choice and design. By selecting a variable to gather data, the boundaries of our possible conclusions are set by the variable selected, the classification categories employed, and the relationships among those categories (nominal, ordinal or scalar).[2] For example, in conducting hazard identification, the possible conclusions about causation are constrained not only by the choice of endpoint to study, but also by the choice of how to measure or classify those adverse effects. The conclusions will be surrounded by

1 For a general discussion of these five types of uncertainty, *See* Vern R. Walker, *The Siren Songs of Science: Toward a Taxonomy of Scientific Uncertainty for Decisionmakers,* 23 Conn. L.Rev. 567 (1991).

2 A "nominal" variable is qualitative only, and yields nominal or categorical data. Its classification categories are merely different from each other, without ranking or ordering among the categories. An example is the variable {color: red/yellow/...}.

An "ordinal" variable is the simplest kind of quantitative variable. Its classification categories have a rank or order, usually reflecting a relative increase in some property, but there is no unit of measurement for the amount of difference between categories. An example is the variable {being hazardous: low/moderate/high}.

Finally, a "scalar" variable is fully quantitative. Its classification categories are not only rank ordered, but they are related by a unit of measurement for incremental frequency, degree, or amount of some relevant property. An example is the variable {length: 1 inch, 2 inches, ...}.

See, e.g., E. Ghiselli et al., Measurement Theory for the Behavioral Sciences, 12-24 (1981).

uncertainty by excluding from our selection of variables.[3] What would we find by studying psychological or pharmacokinetic variables, or by using cardinal instead of nominal measures for those variables? Also, some conceptual inertia occurs once we decide to think about phenomena using selected concepts. Psychological, theoretical and practical influences make us tend to use variables used before.

Concept uncertainty, so understood, seems not currently susceptible to quantification. The concept uncertainty that actually underlies our causal information appears not quantifiable or measurable in any generally accepted way. There is no metric for characterizing how much concept uncertainty is inherent in our causal conclusions. As we will see, however, other types of uncertainty are currently quantifiable in certain ways.

Measurement Uncertainty. Once variables have been selected, a second kind of uncertainty inheres in any data set gathered by using those variables. Uncertainty arises as to the reliability and validity of that data. A measurement method is "reliable" to the extent that it produces consistent results when repeatedly applied to the same individuals.[4] The degree of random scatter within such repeated measurements is the degree of "precision" in the method.[5] An everyday example is the clustered but varying results obtained by weighing ourselves repeatedly on an inexpensive bathroom scale. Unlike reliability, a measurement method is "valid" to the extent that it truly measures what we think it measures.[6] For example, a bathroom scale that produces systemically too high values produces biased data relative to the true weight. Validity concerns the "accuracy" of the measurement data, not its precision.[7]

Measurement reliability is sometimes quantifiable. We can use descriptive statistics such as variance and standard deviation to measure

---

3 One result of such exclusion is overlooking additional risk factors or adverse effects. Another possible result, however, is that causal connections will be misidentified for those endpoints that are studied. *See Causal Uncertainty infra.*

4 Ghiselli et al., *supra* note 2, at 184, 191.

5 *See, e.g.,* Peters & Westgard, *Evaluation of Methods,* in Textbook of Clinical Chemistry 412 (N. Tietz ed. 1986); Mandel, *Accuracy and Prediction: Evaluation and Interpretation of Analytical Results* in 1 Treatise on Analytical Chemistry 256-60 (I. Kolthoff & P. Elving eds. 1978).

6 E. Carmines & R. Zeller, Reliability and Validity Assessment 12 (1979).

7 Peters & Westgard, *supra* note 5, at 412-13.

reliability — the degree of precision or random scatter in repeated measurements.[8] That is, there are ways to measure the degree of inconsistency in repeat measurements on the bathroom scale. Quantifying degrees of validity, however, is far more problematic. By having a criterion or standard method to compare our results, such as a very accurate hospital scale, then we should be able to quantify the extent to which one method (bathroom scale) produces results that agree with results from the criterion method (hospital scale).[9] Characterizations of accuracy other than criterion validity, however, seem to be only qualitative in nature. Instruments to test or measure intelligence, for example, are notorious for the controversy over what they actually measure. Although reliability is generally quantifiable, validity is not, with the possible exception of criterion validity.

Sampling Uncertainty. A third kind of uncertainty is in sampling, which arises once we generalize from sample data to conclusions about a population. Yet, such generalization is the norm in science. Measurements are taken on a sample group, and the summary measures describing that sample data are called "statistics."[10] A population, however, is a larger group from which a sample is drawn. The summary measures for the data that would result from measuring every member of the population are called "parameters."[11] Under certain conditions, current sampling theory enables us to characterize the extent of uncertainty. If our sample is drawn in a random way that allows the computation of probabilities for drawing samples of a given size, then, through significance testing or confidence intervals, we can reach conclusions about whether the (true) parameter is equal to a hypothetical value or is likely inside a given interval.[12] For example, a laboratory study obtaining positive results in animal samples might

8 We can determine the extent to which a set of data exhibits a central tendency, dispersion and form consistent with random variation. For a general discussion of these and other statistical concepts, *See* David H. Kaye & David A. Freeman, *Reference Guide on Statistics* in Federal Judicial Center, Reference Manual on Scientific Evidence (1994).

9 Peters & Westgard, *supra* note 5, at 412; Carmines & Zeller, *supra* note 6, at 19.

10 H. Loether & D. McTavish, Descriptive and Inferential Statistics: An Introduction 4-5 (2d ed. 1980); W. Hays, Statistics 190-92 (4th ed. 1988).

11 *See* Loether & McTavish, *supra* note 10, at 6.

12 Walker, *supra* note 1, at 590-98.

warrant that there is probably a real increased risk due to exposure (because "the sample results are statistically significant") or that the true risk probably falls within a specified range (the "confidence interval"). We can also compute the "power" of the study to detect whether the (true) parameter differs from a hypothetical value.[13] These three techniques of inferential statistics (significance testing, confidence intervals and power analysis) are means of characterizing the degree of sampling uncertainty inherent in conclusions about populations.

Thus, sampling uncertainty is quantifiable if the selection is sufficiently random that a probability can be computed for samples. Such sampling is still no guarantee that our predictions about the population value will be correct: sampling uncertainty is inherently probabilistic whenever it occurs. If the sampling protocol does not yield a probability sample, however, it is uncertain whether the drawn sample is actually representative of the entire population. Without a probability sample our sampling uncertainty is generally unquantifiable.

Mathematical Modeling Uncertainty. A fourth kind of uncertainty arises when we use models to predict the values of one variable from the values of other variables. Whereas concept, measurement and sampling uncertainties occur even when we are dealing with data from only one variable, modeling uncertainty occurs when the predictive relationships among variables are not as simple as our model assumes. Examples of predictive mathematical models are regression analysis and relative risk.[14] Using either type, we risk error when using exposure data, for example, to predict the disease's incidence in an exposed group.

With some mathematical models we can quantify to some extent the expected degree of predictive error. In linear bivariate regression

---

13 Technically, the "power" of a statistical test is "the probability of being right in rejecting H_0 [the hypothetical value] given that H_i is true," where H_0 and H_i are possible values for the parameter in the population. Hays, *supra* note 10, at 248, 263-64.

14 Regression analysis is a specific way of analyzing a variable to be predicted (the dependent variable) using one or more predictor variables (independent variables). *See, e.g.,* Daniel L. Rubinfeld, *Reference Guide on Multiple Regression* in Federal Judicial Center, Reference Manual on Scientific Evidence, *supra* note 8, at 415.

Relative risk is the ratio of the incidence of an adverse effect in a group of individuals with a risk factor (e.g., exposure) relative to the incidence in a group similarly situated except for that risk factor. *See, e.g.,* Linda A. Bailey et al., *Reference Guide on Epidemiology* in Federal Judicial Center, Reference Manual on Scientific Evidence *supra* note 8, at 121, 147-49.

models, for example, the coefficient of determination (r2) and Pearson's correlation coefficient (r) are measures of predictive uncertainty because they quantify the strength of the linear association between the predictor and predicted variables.[15] Yet, other kinds of modeling uncertainty are less quantifiable — such as uncertainty due to the form of the selected mathematical equation. To determine the amount of uncertainty that is due to using a linear model over a nonlinear model, for example, then we may have to fit both forms to the data and conduct a sensitivity analysis between them.

Causal Uncertainty. The fifth kind of uncertainty inherent in causal information about groups is causal uncertainty itself. This is created by making an inference from statistical associations established by mathematical modeling to conclusions about causation. While mathematical models and statistical associations can sometimes provide decent predictions, only causal models provide true explanations of why *what we observe* happens as it does. Causal explanations depend upon what risk assessors call "hazard identification"[16] and what courts sometimes call "generic causation."[17] Generic causal capacity is the agent's capability to cause some type of effect, even if not always. An example is whether a certain chemical can cause human cancer.

Statistical association does not entail generic causation. Even finding a statistically significant association between exposure and illness in a study, does not mean that the exposure is necessarily a causal factor in causing the illness. For example, an unstudied variable may cause the observed association that is related both to exposure and to illness.[18] Moreover, a lack of statistical association between variables is not sufficient evidence of a lack of causation. For example, an antagonistic action between two study variables could tend to mask the causal relations among the variables.[19] While statistical association provides

15 Walker, *supra* note 1, at 604-05; *See* Hays, *supra* note 10, at 554-60.

16 National Research Council, Risk Assessment in the Fed. Gov't.: Managing the Process 19-23 (1983); National Research Council, Science and Judgment in Risk Assessment 26, 57-60 (1994).

17 *E.g.,* Sterling v. Velsicol Chem. Corp., 855 F.2d 1188, 1200 (6th Cir. 1988).

18 *See, e.g.,* J. Davis, The Logic of Casual Order 25-27 (1985).

19 *Id.* at 33.

There may also be a real association in the population but the study does not have sufficient statistical power to detect it. This would not be a problem of causal

our best evidence of causation, an inferential step is involved in reaching causal conclusions on the basis of statistical evidence.

It is not easy, and may be impossible, to quantify the residual causal uncertainty created by this inferential step, or the extent of the reduced possibility of causal error in our conclusions. The adequacy of a causal account remains largely a professional judgment matter. It requires the evaluation and combination of many factors that we cannot quantify, such as the plausibility of a causal mechanism posited by a theory.

Information about Specific Individuals

In addition to the above five types of uncertainty inherent in scientific conclusions about causation in groups, there are two additional areas of uncertainty associated with conclusions about specific individuals (as opposed to groups).

Direct Inference. First, there are uncertainties inherent in a "direct inference" from statistical information about a group to a probabilistic conclusion about any specific individual in the group. Suppose scientific research supports a causal model that relates a certain pathway and level of exposure to an increase in the illness's incidence. Suppose further that the data show biological variability in the human response to this exposure. That is, some people do and do not develop the illness. Some who develop the illness develop it more quickly than others or with more severity. We are often interested in using such group information to make a prediction about a specific individual: if this specific individual were exposed in that way, what is the probability that he or she would develop the illness? Sometimes we are interested in providing a retrospective causal explanation: if this specific individual has the illness and has been exposed in that way, what is the probability that the exposure caused the individual's illness? (The courts often require plaintiffs to prove such "specific causation.")[20] The predictive question is often asked to guide a course of action (such as taking a medication or voting for an incinerator) and the question about explanation might be asked in deciding compensation (as in a toxic tort lawsuit). Given

uncertainty. however, but one of sampling uncertainty. The process of sampling created the misleading lack of association in the sample, although the lack of association in the sample led to mistaken conclusions about generic causation.

20 *E.g., Sterling,* 855 F.2d at 1200.

the variability in the human response to the exposure, we can interpret the questions as inquiries about where the specific individual falls within that group variability. Are we warranted in using the group statistics to directly infer a probabilistic conclusion about a specific individual? For example, if 35% of exposed individuals are expected to develop the illness, is there a 0.35 probability that a specific individual will do so?

There are new sources of uncertainty inherent to such a direct inference — in addition to already discussed uncertainties inherent in group information from which the inference begins. These stem from two considerations. First, for any variable, the measurement "score" for the specific individual is either known or unknown. Second, the generic causal relevance of that variable to the illness is either adequately understood or not. These two considerations form a matrix with four combinations, as shown below.

Figure 1
For any Variable

Specific Individual's Value or Score:	*Causal Relevance Understood*	
	Yes	No
Known	A	B
Unknown	C	D

Each of these combinations is typically present in direct inference about risk, due to the great number of possible variables. Moreover, each adds a degree of uncertainty to any conclusion about what is likely to happen (or is likely to have happened) in specific individual case. The four combinations or "cells" shown in Figure 1 are:

A. Information is "known" about the specific individual and the generic causal relevance of that information has been adequately modeled. We may know, e.g., that a certain level of exposure is a causal risk factor and that this specific individual experienced it. This combination makes the uncertainties inherent in the generic causal model relevant and raises the possibility that the individual's exposure was not accurately described.

B. Information is "known" about an individual but its causal relevance is unknown. For example, we may know that s/he has a family history of emphysema, but we do not know whether that history is a risk factor for the concerned illness.

C. The generic causal relevance of some risk factor has been adequately determined in scientific group studies, but

we do not know an individual's score on that factor. We may, e.g., have good reason to believe that a family history of high blood pressure is a risk factor, but we do not know whether this individual has such a family history.

D. Some facts about the specific individual are unknown and their causal relevance is also unknown. This is true about most of the genetic makeup of an individual, as well as about events in the individual's developmental and environmental histories.

Each combination represents uncertainties in any prediction or explanation about specific cases. The etiology of few illnesses is so completely understood that we can predict specific case outcomes.

Some of the uncertainty in the group information relevant to the first and third combinations is quantifiable, as discussed earlier in this article, but the remaining uncertainties are largely unmeasurable. The latter derive from questions of fact whose answers are unknown, and the uncertainty inherent in any answer is unmeasurable.

Specific Classification. A second kind of uncertainty inherent in information about specific individuals is that in classifying them under a variable. It is a kind of individual "measurement uncertainty" and is the primary alternative to concluding on direct inference. When classifying a specific individual under the categories of a variable, such as "this person suffers from Grade 1 byssinosis," we justify it by explaining the relevant variable ("having byssinosis") and the criteria for classification, together with any relevant perception-based observations about the individual. But a possibility of error in conducting the individual measurements is partially reflected in the reliability and validity data on the measurement technique itself. Scientific studies begin with selecting a variable, assembling reliability and validity data on the measurement instrument, gathering data using it, and then analyze resulting group data. We also use such group results to warrant specific case classifications; our warrant for this includes explaining the reliability and validity of the particular measurement process involved and others, e.g., the process of perception itself. When questioned about classifying a specific worker as having Grade 1 byssinosis, the classification's justification is based on our perceptions (and their reliability and validity), and on the reliability and validity of other measurements and instruments. This also happens when researchers

encounter an extreme outlier or highly anomalous result: They resort to specific causation to explain the specific case. Measurement data are certainly foundational for science. But each individual measurement is also always open to inquiry as to its truth, and each measurement or classification has inherent its own peculiar uncertainties. Some of these are to some extent quantifiable; others are not.

Conclusion

Even this brief survey shows numerous types of inherent uncertainty, many unquantifiable. They are logically distinct and cumulative, but we do not know how to quantify their cumulative effect. One movement within risk assessment is to deal with uncertainties by turning away from objective probability determinations about the relative frequency of events and toward the measurement of subjective probability — the subjective degree of confidence about a prediction or explanation. Resort to subjective probability partly attempts to measure all uncertainties in an integrated way, not by measuring the relative frequency of the observable events or subevents, but rather by measuring the degree of an expert's subjective confidence.[21] This also allows us to assign probability functions to descriptions of unique events and information about specific individuals — such as the probability that a specific individual did or will develop cancer from a specific exposure event.[22] The degree of subjective confidence is (at least in theory) a measurable phenomenon, even if it is not what we set out to study.[23] Yet, it is critical that determining the subjective confidence of a specific individual about any proposition's truth has itself tremendous inherent uncertainties. It has all the those inherent in classifying a specific individual scientist under a variable called "degree of confidence," as well as those inherent in deciding

[21] I have elsewhere discussed the problem of "epistemic uncertainty". Such uncertainty derives from the differences in our current ways of understanding the concepts fundamental to all causal conclusions — the concepts of deductive logic, probability theory, and mathematics. *See* Walker, *supra* note 1, at 618-24.

[22] For a general discussion of the logical problems associated with direct inference and probabilistic propositions about unique events or specific individuals; *see* Vern R. Walker, *Direct Inference in the Lost Chance Cases: Factfinding Constraints under Minimal Fairness to Parties,* 23 Hofstra L. Rev. 247 (1994).

[23] We set out to study the exposure effects but end up studying the scientist who studies the exposure effects.

whether that specific scientist's degree of confidence is a reliable and valid indicator of observable phenomena. We should try to determine precisely what the subjectivist "solution" to uncertainty accomplishes, other than contributing additional variables and uncertainties to an already complex problem.

Within the five kinds of group uncertainty and the two kinds of specific uncertainty, some elements are quantifiable, but many are not. Quantification usually consists of isolating logically distinct components of the reasoning process and devising methods for placing probability distributions on their outcome possibilities. Sampling theory remains a major example of such a technique. Still, even when an element of uncertainty is quantifiable, it remains uncertainty. Our conclusions might be in error, even though supported by reasoning that sugggests a high probability of their being true.

If the above taxonomy and analysis of uncertainty is correct, then even the simplest information about risks, benefits and costs has at its base conclusions about generic and specific causation. Such premises have inherent uncertainties of many different varieties; some can be reduced, but none can be eliminated. Uncertainties will always be inherent, even in theory. Moreover, matters are worse in practice. In all risk regulation cases, decisions are made in real time. Someone must decide whether the information is "good enough," in the sense that the residual uncertainties are theoretically or practically acceptable. Thus, due to both theory and practice we must resort to "decision rules" for when and how to proceed despite these many "faces" of uncertainty. Such decision rules need to be evaluated and justified with an eye toward promoting the optimal blend of effectiveness, efficiency and equity. Risk regulation is, in the end, regulation, and the optimal combination of effectiveness, efficiency and equity is all we can ever hope to achieve. Perhaps we can better approximate that societal ideal, however, if by evaluating decision rules on a precise case-by-case basis, after we clarify the many "faces" of uncertainty.