Protecting People in Spite of—or Thanks to—the “Veil of Ignorance”

ADAM M. FINKEL

When society determines that an action to protect public health or the environment is warranted because its benefits exceed its costs, many of us probably conjure up a mental picture of a balance between two discrete quantities, tipping positive. Presumably, all but the most self-contradictory forms of the precautionary principle (discussed by Resnik earlier in this volume) share this mental construct: that is, that the positives that accompany an action outweigh the harms caused by that action or the harms that inaction would perpetuate, even if no attempt is made to gauge costs and benefits in the traditional sense. But these are both abstractions of cost and benefit, and they result from aggregating health, environmental, and economic effects that apply in disparate ways to every individual. When a cost or benefit outweighs its counterpart, it is because a collection of one outweighs that of the other. By looking at only the total bottom line, we can, and often do, luck into sound choices. An individualized assessment, however, can improve collective action in two fundamental and not mutually exclusive ways: (1) it can allow for fine-tuning of the intervention to provide more
protection to the individuals most in need or less stringency for those who do not need or prefer not to pay for additional protections; and (2) it can change the level of protection we provide across the board, in light of what we learn about the distribution of costs and benefits.

To see both the whole and the sum of its parts in these situations requires both the will to examine individual costs and benefits and appropriate tools for discerning effects on individuals. The main thesis of this chapter is that for the past twenty years or more, our technical capacity to individualize estimates of risk and benefit has increased much faster than has our willingness to make use of these new abilities.

This situation is not simply the result of a lack of vision, as I and others have argued previously (Finkel, 1984; Hattis, 2004). In fact, recent developments suggest that some of the current reticence to exploit our increasing ability to individualize risk stems from legitimate concerns about the side effects of doing so. In this chapter, I attempt to distinguish some of the inevitable downsides of individualized environmental and occupational health protection from other pitfalls that are avoidable. I then sketch out an alternative approach to reconcile the desire to protect individuals according to their unique risks with the reluctance to proceed down such a path: namely, interventions informed by human genetic (and other) variability that are not dependent on identification of the specific individuals motivating society's concern.

This chapter is structured to support the thesis that personal and social decisions that acknowledge variability in risk and cost are more sensible and robust than those that ignore variation or "average it away." In light of the benefits and harms of identifying individuals according to their place on the distribution of risk or cost, I emphasize the concept of "anonymous protection" as a win/win solution to the tension between individualization and identification, or as a worthwhile fallback position that may have unique virtues.

With a debt to John Rawls, who described as the "veil of ignorance" the situation in which individuals know the existence of distributed attributes but not their own personal circumstances (Rawls, 1971), I focus primarily on reasons why this partial "ignorance" may be preferable to complete awareness, despite the connotations of that word. Rawls looks longingly on an unlikely state of ignorance regarding wealth and other obvious attributes, because he sees it as in fact superior to more complete knowledge: "one excludes the knowledge of those contingencies
which sets men at odds and allows them to be guided by their prejudices" (Rawls, 1971, p. 19). In the regulation of environmental and occupational hazards, in contrast, we can choose to place (or to allow) a "veil of ignorance" over less obvious individual knowledge, and in so doing we may be able to promote justice at minimal risk to efficiency.

Before sketching out this possible solution, I offer two prefatory sections. In the second section, I place the attention to individual risk in the context of similar lapses in several related science-policy arenas. I summarize the special concerns raised by individual genetic information relevant to environmental or occupational exposures in the third section. In the fourth section, I discuss some conceptual objections to calculating and acting on information about "nonidentifiable variability," and in the fifth section I present several sets of reasons to explain why such objections may be misguided as well as unduly pessimistic. Finally, I discuss some policy implications of incorporating nonidentifiable variability into risk-based decisions about environmental and other issues.

"The Fault Is in Ourselves:" Mishandling Variation May Outlast Our Ignorance of It

It should come as no surprise that our ability to pinpoint environmental risks to individuals has progressed faster than our success in acting appropriately on such knowledge. There are many analogous situations where we have had requisite knowledge of interindividual variability at hand, and yet we have ignored, misinterpreted, or misapplied that knowledge. Such situations can enshrine a self-fulfilling supply-and-demand problem, as failing to appreciate the benefits of understanding variability in one area can make understanding variability in other areas seem worthless or even counterproductive. The following examples are meant to shed light on two different types of failures in responding to available information about variation. They emphasize how the task becomes even more difficult in those other cases where we need both to generate and to incorporate data on interindividual differences.

Ignoring Variability

Sometimes we utterly fail to acknowledge variation that is unambiguously present. Several examples from different contexts are provided below.
Environmental Exposures

Goldstein et al. (1992) analyzed the fatality risks from airplane crashes to persons not on board the plane ("groundlings"). They tallied one hundred fifty such deaths in the United States over an eleven-year period (1975–85), divided by the size of the U.S. population in 1980, and converted from an eleven-year window to a seventy-year lifetime, to arrive at an individual lifetime risk of $4.2 \times 10^{-6}$. Nowhere in this short paper do the authors even hint at interindividual variability (due primarily to location) in risk; rather, the authors refer to "the risk" as a precisely known quantity and define it as "the likelihood of death due to an airplane crashing on you while you are on the ground" (emphasis added).

A conclusion drawn from this exercise—that the risk of being hit by a falling airplane is roughly 4.2 times greater than the one-per-million level of risk that sometimes prompts governmental action—later became a central focus of congressional debate (1993–98) over a series of legislative "regulatory reform" proposals to change the way agencies assess and manage risks. John Graham testified on various occasions during this period (Graham, 1993; Graham, 1994) that Congress might rethink the $10^{-6}$ benchmark, urging members of Congress to consider whether society should strive to reduce risks to a level four times smaller than the implicitly trivial risk each baby born in the United States faces from "being killed while on the ground by a crashing airplane" (Crain, 1993).

Around that time, I suggested (Finkel, 1996) that "these average risks do not reflect the substantial variance in risk that real people undergo: the reason most people probably think the risk of being hit by a falling airplane is remotely low is that for most of us it is remotely low." I then offered a "guestimate" that the $4.2 \times 10^{-6}$ population average masked much smaller risk (I guessed $10^{-7}$) to perhaps 99 percent of the population living far from any airport flight path, and a much larger risk (I guessed $1.2 \times 10^{-5}$) to the remaining 1 percent of Americans living at or near the boundary of airport property.

Five years later, Thompson, Reboiu, and Cooke (2001) analyzed each of the one hundred fifty fatalities and showed that about 3 percent of U.S. residents live less than two miles from an airport runway and face a lifetime "groundling" risk of approximately $1.2 \times 10^{-4}$, while for the remaining 97 percent of us, this risk is below $6 \times 10^{-7}$. Advancing the original
analysis by one step, from the assumption of no variability to a dichoto-
mous assumption, reveals that for some of us the "grounding" risk is at
least two hundred times larger than it is for others. Qualitatively, of course,
the two subpopulations' risks lie on opposite sides of the $10^{-6}$ reference
point, suggesting that it may have been particularly misleading to average
a trivial risk to nearly everyone with a large risk to a small minority and
not to disclose or notice the implicit calculation. Note that in this example
proximity to an airport is something observable, without the need for any
sophisticated analysis of an individual's characteristics. This points to the
allure (for the analyst, and perhaps for the recipients of such "risk in
perspective") of reducing dichotomies or continuous variation to a single av-
erage, even in cases such as these where the variability is so tangible.4

Susceptibility to Environmental Disease

Cancer risk-assessment methods depend on many assumptions that
allow us to estimate risks in the face of incomplete data and imperfect
models. Whenever we need to translate information found under the
proverbial lamppost (where we can observe some phenomenon) to a sit-
uation of greater interest where no direct observations are available, we
must either assume equivalence or difference, two categories that are mu-
tually exclusive and exhaustive. For example, because we know that mice
and humans breathe different quantities of air during a given time pe-
riod, we extrapolate across species by trying to quantify that difference.

Extrapolation within a single species, whether this is recognized or
not, similarly challenges us to quantify differences or else to assume that
no meaningful differences exist. Risk assessment for noncarcinogenic
compounds has relied heavily for several decades on quantitative infer-
ences about human interindividual variability. The U.S. Environmental
Protection Agency (EPA) and other agencies routinely decrease allowable
exposures by a factor of ten to account for the possibility that some hu-
mans are more susceptible than the typical human, who presumably
would be protected appropriately at the higher level of exposure. A sub-
stantial literature suggests that, depending on the compound and assump-
tions involved, the factor of ten either may be needlessly large (Dourson,
Felter, and Robinson 1996) or that it may be insufficient to protect indi-
viduals who are substantially more susceptible than the typical person
(Hattis, Banati, and Goble, 1999). However, the principle of trying to
admit and accommodate intraspecies differences in noncancer risk
assessment will surely survive the ongoing fine-tuning of the precise amount of adjustment.

The situation is completely different with respect to cancer risk assessment, however. EPA and other agencies have, for decades, implicitly treated all humans as identical in susceptibility to the carcinogenic effects of environmental exposures, despite a wealth of information suggesting otherwise. The information that human interindividual variation is substantial began to accumulate before EPA's initial codification of its carcinogenic risk assessment procedures in the late 1970s and is accumulating currently (Finkel, 1995a; Hattis and Barlow, 1996). Nonetheless, whenever cancer risk assessors estimate human risk based on animal bioassay information, they derive a dose-response relationship that is valid for only one hypothetical human being, but they apply the relationship to all persons. In theory, when extrapolating from laboratory animals to humans, it is appropriate to assume that the susceptibility of the test animals (essentially a point estimate because they are inbred for homogeneity) is equivalent to that of the median human (National Research Council [NRC], 1994, p. 210; Hattis and Goble, 2003).

This problem also pervades risk estimation that relies on human epidemiologic data, but here the single, nonvarying dose-response relationship can reflect only the average susceptibility of the human population studied. But this average can differ both at random or systematically from the average susceptibility of the broader population for whom risk is being estimated (Finkel, 1995a), as well as from any individual member of that population. In its 2005 Guidelines for Carcinogen Risk Assessment (p. A-9), EPA does make one passing reference to the "range of human variation" in susceptibility, but only to make the comforting claim that its practice of using a linear (rather than, presumably, a threshold or sub-linear) dose-response function "adequately accounts for human variation" (U.S. Environmental Protection Agency [EPA], 2005a). EPA is apparently asserting that the "conservatism" it believes to be inherent in the linear model makes up for concern about individuals of higher-than-average susceptibility—in effect, asserting that it overestimates risk for everyone but just does so less dramatically for some people. This claim is unsupported in that no data exist that suggest that the linear model is in fact conservative or that any upward bias that might occur at this step is sufficient to make up for the downward bias inherent in ignoring the half of the population with above-median susceptibility. Although at least
EPA now admits that humans (unlike laboratory animals bred for genetic homogeneity) have individual dose-response relationships that vary across the population, it does not take this epiphany into account in its cancer risk assessments (with one exception discussed below), as if human variability was a theoretical nicety, rather than a profound challenge to the validity and policy relevance of the agency’s cancer risk assessment outputs (NRC, 1994; see especially the committee’s recommendation on p. 219).

Medical Decision Making

In contrast to the relationship between the environmental risk manager and the public affected by these risks, the relationship between the clinician and the patient is inherently personal, with the unit of analysis (at least ostensibly) being the individual, rather than the group or population. Physicians are expected to be sensitive to obvious characteristics of the patient that individualize the diagnostic process; for example, the differential diagnosis of hearing loss should not proceed identically in two successive patients, one a teenager and one an octogenarian. To my knowledge, however, no systematic study has ever been carried out to determine the extent to which clinicians successfully (or even attempt to) individualize diagnosis and treatment when such decisions hinge on quantitative variation in risks, even when the underlying variation is obvious. In my experience, supported by many discussions with colleagues who have either given or received medical advice based on quantitative information, physicians often apply population-average values of key probabilities to patients regardless of the extent to which the patient’s own probability would seem likely to differ from the average. (See, for example, a book written largely about this subject, Schneider and Lane, 2005.)

As an example, my wife and I faced on several occasions the happy but highly nerve-wracking choice of how many embryos to transfer during in vitro fertilization cycles at several different clinics. In every case, when we expressed concern both about failing to conceive and about the risk of twins and higher-order multiple births, our doctors cited the results of the same recent large trials that had established the average probability (p) of successful implantation per seven-day blastocyst transferred. We became adept at rapidly doing simple binomial expansions to allocate the probabilities of 0 through n births as a function of p and the
number of blastocysts we might choose to transfer, invariably (pun intended), when we asked whether we should make a decision based on the population value of \( p \) or on some individualized value that might be higher or lower based on our own characteristics (age, reproductive history, previous attempts, etc.) and on the morphology of the actual blastocyst involved, we were told that the studies had provided all the information available—one value of \( p \) for all women in my wife’s five-year age category. However, our decision was rather sensitive to small changes in \( p \) (Table 17.1). At the population \( p \) of 0.5 per blastocyst, we likely would have transferred only one blastocyst to avoid the 25 percent chance of twins with transfer of two. On the other hand, if our individualized \( p \) was equal to 0.4 (and we believed that our \( p \) was lower than 0.5, due to factors our doctors were well aware of), we would have viewed transferring two as preferable to one, as transferring two would have resulted in a much lower probability of failure with only a 16 percent chance of twins.

My suspicion that clinicians commonly provide quantitative information unmodified by consideration of likely (that is, neither speculative nor difficult-to-discern) factors specific to the patient(s) receiving it is bolstered by a preliminary examination of the medical decision-making literature. Only relatively recently (Winkler and Smith, 2004) have researchers begun to consider methods for communicating uncertainty in one of the most widely used probabilistic analyses—the Bayesian posttest

<table>
<thead>
<tr>
<th>Number transferred</th>
<th>Number of live births</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>.6*</td>
</tr>
<tr>
<td></td>
<td>.51</td>
</tr>
<tr>
<td></td>
<td>.42</td>
</tr>
<tr>
<td>2</td>
<td>.36*</td>
</tr>
<tr>
<td></td>
<td>.25*</td>
</tr>
<tr>
<td></td>
<td>.16*</td>
</tr>
<tr>
<td>3</td>
<td>.215*</td>
</tr>
<tr>
<td></td>
<td>.125*</td>
</tr>
<tr>
<td></td>
<td>.08*</td>
</tr>
</tbody>
</table>

Notes: dash = not applicable.

*p = .4
*p = .5
*p = .6
probability of disease given a medical test with an assumed sensitivity, specificity, and population prevalence (pretest probability). These articles seem to concentrate on how true uncertainty in one or more of these three parameters affects the patient’s posttest probability, rather than the situation where the population parameters can be precisely estimated, but simply differ from patient to patient. One would expect, for instance, the prior disease probability to vary from person to person, based on age, related symptoms, and many other factors.

The tendency to “average away” known or readily ascertainable variability in the clinical setting appears to be writ large in the application of decision theory to public health. Hundreds of analyses of the cost-effectiveness of particular surgical interventions, pharmaceutical therapies, and diagnostic tests have been published over the past twenty-five years, and most of them report either a single value or a small number of subgroup values for the cost-effectiveness result (e.g., a given test may be “cost-effective” for men and not for women, or for women of childbearing age and not for others). These population cost-effectiveness numbers wield tremendous influence in determining the standard of care for particular conditions, and in determining which interventions will be covered by insurance, and if so, at what frequency. However, any estimate of the cost of a procedure per life-year extended (or a similar metric gauging mortality or morbidity reduction) will vary over the population to whom it applies, in proportion to individual differences in baseline risk, efficacy, valuation, and other factors. Although recent articles in the medical decision-making literature (Zarin, 2003) have begun to explore biases and uncertainties in population cost-effectiveness ratios resulting from ignoring interindividual variability, relatively few sensitivity analyses have been undertaken to explore which decisions might be optimal for populations but suboptimal or flatly incorrect for individuals within them who differ from the group.

Economic Welfare

If risk scientists’ progress in acknowledging and quantifying known interindividual variation in risk has been rudimentary, the situation with respect to variation on the cost side of the cost-benefit ledger is even less well-developed. Regulatory economists often define “costs” as the resources that those who must comply with a regulatory or other intervention must expend. A growing literature (see, for example, Goodstein and
Hodges, 1997; Harrington, Morgenstern, and Nelson, 2000) explores the biases, more often than not upward ones, surrounding current estimates of total compliance cost. More recent articles (Pizer and Kopp, 2003) make the point that even an unbiased estimate of total compliance cost is a surrogate for the correct measure, which is social cost—the sum total of economic changes, both negative and positive. These economic changes include compliance costs, but also price changes and benefits and costs to workers and firms that supply the controls or knowledge that the initial complier must purchase (Porter and van der Linde, 1995; Berck and Hoffman, 2002). But even an accurate estimate of total social cost is itself only a measure that is insensitive to the distribution of individual costs, just as a measure of the number of lives a regulation is expected to “save” is insensitive to the distribution of risk-reducing benefits across individuals. As Pizer and Kopp report, “given the pervasiveness and magnitude of environmental regulation, one would think that comprehensive studies of the cost and benefit distribution of these policies would be bountiful. Ironically, the contrary is true” (p. 33).

As a result, when decision makers compare total benefits with total costs (C), the individual citizen (in a population of size P) has no basis for knowing whether his own share of those total costs will be near zero (or even less than zero), near C, or anywhere in between. The assumption that costs are always borne equally across the population (i.e., that everyone’s share is exactly C divided by P) may be useful for some of the population in some cases, but mainstream regulatory economics gives no hint that this may be the rare exception, just as in the “groundlings” example above, where the putative benefits rather than the costs of control are concentrated, but are not acknowledged as such.

Both risks and costs vary across individuals, so the distribution of their ratio or their difference will generally be even broader than the distribution of risk or cost viewed in isolation (Finkel, 1995b). A rejected intervention for which “the (total) costs exceed the benefits,” therefore, may mask an underlying reality in which the individual benefit would exceed the individual cost for the majority of affected citizens or vice versa. This is a mathematical property of distributed quantities and does not hinge on the colloquial definition of equity—giving special weight to individuals who are disproportionately affected. Even giving each person identical weight can cause us to conclude that the typical value for individual net benefit may be of a different sign than the value of aggregate net benefit.
These examples show how tempting it often is to regard interindividual variability as an “annoying detail” (Hattis, 2004) that can safely be averaged away, as qualitatively less important than reaching conclusions about group behavior or between-group differences, or even as an impediment to studying those very differences. In risk assessment for environmental health, any impetus to explore and respond to interindividual differences may wane simply because alternatives to expressing risk quantitatively are on the rise. Both the precautionary principle (touted largely from outside the federal regulatory system) and outputs of analysis such as the reference dose or the “margin of exposure” that do not use risk or harm as their currency (developed largely from within the agencies) can drive decision making that is insensitive to risk and hence by definition uninterested in within-population variations in risk. 10

Mishandling Variability

Even when we acknowledge or emphasize human variation, examples abound of misusing or oversimplifying the information added. I offer six examples of different ways in which disaggregating the population into two or more subgroups can cloud or hinder communication or intervention. These examples are biased toward recent events and commentaries, and for this reason may not be the most apt instances of the phenomena I am trying to categorize. However, they are meant to stimulate thought about the wide array of arenas in which we end up taking initial but unsatisfactory steps toward individualizing risk or benefit.

1. Splitting the population via a characteristic not causally related to risk, or only partly correlated with it. The U.S. Food and Drug Administration (FDA, 2004b) published draft guidance that various facilities can use to determine, among other things, which prospective donors of cells and body tissues should be deemed ineligible due to the possibility of HIV transmission to the recipient (during the several-month period in which HIV testing of the donated material may yield a false negative). In particular, the guidance suggests that sperm banks and fertility clinics should reject sperm donors who have had homosexual sex during the preceding five years. However, FDA also recommended that a prospective donor who has had heterosexual sex with a known HIV-positive person more than twelve months before donation be deemed eligible to
donate. Many activists and clinics objected to these recommendations on the grounds that they attach a greater stigma to less risky (monogamous, “safe,” possible or documented HIV-negative partner) homosexual practices than they do to riskier (polygamous, unprotected, possible HIV-positive partner or partners) heterosexual practices. FDA acknowledged (FDA, 2004a, pp. 29805–6) that this categorization might incorrectly exclude many prospective donors whose risk of transmitting HIV was lower than other included donors, but noted that one of its advisory committees had determined there were no data that could identify “subsets of men who have had sex with other men . . . [whose risks are] similar to the population at large.”11 Essentially, there are at least two ways to dichotomize a population for such a purpose: via risky behavior or via a different behavior that is imperfectly correlated with risk. FDA has chosen the latter course, putting the onus to improve the sorting mechanism on those who would remedy the new inequities.12

2. Splitting the population via a relevant variable, while ignoring a more powerful one. Many of the other authors in this volume have referred to the “genetics loads the gun, environment pulls the trigger” model of causation, which holds that few cases of chronic disease arise without some influence from both predisposition and exposure. A possible corollary to this general rule would be that both genetic and environmental variability must be considered when trying to explain the presence or absence of disease in an exposed population. One tidbit of recent anecdotal information may call to mind the frequency of inferences that go astray due to only partial acknowledgment of interindividual variability. A letter to the New York Times (Schlack, 2005) notes that “boys account for 80% of all autism cases. Are we to believe that boys received more of the thimerosal vaccine than girls?” Without reaching a conclusion about the existence of a relationship between thimerosal exposure and autism, surely an explanation exists for the gender difference that requires no systematic difference in exposure patterns: boys may simply be more susceptible to the adverse effects of a given exposure to thimerosal than girls are.

3. Splitting the population via relevant, powerful characteristics that apply to groups rather than (necessarily) to the individuals within them. The well-known “ecologic fallacy” can yield misleading conclusions about causality or the quantitative strength of a correctly inferred causal relationship, because comparisons of group characteristics may mask con-
flicting or attenuating information about the characteristics and outcomes of individuals. Even in the presence of a correct and precise inference about group behavior, however, policies applied to individuals may be inequitable. One of the “new ideas of 2005” featured in an annual review in the New York Times Sunday Magazine—recommending criminals for shorter or longer sentences based on demographic characteristics—may epitomize the dilemma of policies that may be appropriate for groups but not for individuals. Reporter Emily Bazelon (2005) characterized the use of demographic and behavioral risk data by Virginia and other state sentencing commissions as “beginning to make it possible to determine which bad guys really will commit new offenses... based on a short list of factors with a proven relationship to future risk.”13 She reported approvingly that when the algorithm was tested on prisoners who had already served their sentences, 12 percent of the felons who would have scored at or below the thirty-five-point cutoff (and therefore would have been recommended for probation or house arrest) committed new crimes, while those who would have scored above thirty-five points had a recidivism rate more than threefold higher (38%). Various critics of this kind of system have focused on the disconnect between most of the attributes in the risk algorithm and “blameworthiness”—some of the variables being immutable (e.g., age) and others (e.g., marital status) reflecting individual opportunities as well as preferences. Bazelon noted that “a woman in her 40’s who deals drugs hasn’t done anything more to earn trust or deserve a break than a male dealer in his 20’s charged with the same offense.” But this sort of criticism focuses on the broad relationship between the variables and the outcome, rather than also considering the statistical relationship between the individual and the group “dose-response” functions. The 3:1 relative risk separating those with high and low scores is impressive, but it also implies that 62 percent of those at “high risk” turned out to have been better candidates for mercy than 12 percent of those at “low risk,” even though the all of the former offenders would have received harsher sentences. To be sure, the real question here is what the sum of type I and type II errors would be (possibly higher) had the sentencing commissions considered none of the demographic factors and assigned the same sentence to everyone committing a particular crime. This counterfactual gets to the heart of “fairness” in a situation in which we can reduce the misclassification rate by actively tinkering with the intervention or can choose to treat everyone identi-
cally. By the former stratagem, some people will receive punishments doubly inappropriate for their true but unknown individual risk of recidivism. The latter approach will probably create more prevalent albeit less severe inequities.

4. Splitting the population via a relevant explanatory variable, but one for which the within-group variance may exceed the between-group variance. With the publication of its 2005 "Guidelines for Carcinogen Risk Assessment," EPA took the first small step toward "individualizing" its cancer risk assessments to account for each person's unique susceptibility to carcinogenesis. It divided the human population into three age groups and mandated an additional upward adjustment in some cancer potency factors for two of these groups. For carcinogens thought to have linear dose-response functions in humans, the guidelines incorporate a tenfold upward adjustment for exposures that occur between birth and two years of age and a threefold adjustment for exposures that occur between ages two and fifteen (EPA, 2005b).34

According to EPA, therefore, infants as a group are ten times more susceptible than are adults, and exposures during this life stage would need to be one-tenth as high as those allowable for any two-year period in adulthood in order to maintain an acceptably low probability of cancer in either group. The magnitude of this adjustment, however, comes from analysis of many bioassays in which the apparent carcinogenic potency of a given substance in juvenile versus adult animals could be compared (with a "supporting role" given to the one relevant human database, the comparison of cancer incidence in Japanese exposed to atomic bomb radiation as children versus those who were adults in 1945). Because residents and other test animals are inbred to minimize genetic heterogeneity and are kept in controlled environments that minimize the influence of factors such as disease and concomitant exposures on susceptibility, any observed differences in potency between juvenile and adult animals reflect between-group variations exclusively. However, epidemiologic and biochemical studies of interindividual variability within subgroups of the human population (Finkel, 1995a; Hattis and Barlow, 1996), suggest as a first approximation that "typical" adults (and typical children) can differ from each other in their overall susceptibility to carcinogenesis by a factor of twenty-five to fifty or even more, due to differences in metabolism, DNA repair, immune surveillance, and similar factors. An individual adult may therefore need the extra "conservatism" in potency
estimation far more than does the average or the “resistant” infant, but only the infants will benefit from EPA’s new foray into assessing cancer risk to some individuals.

5. **Splitting the population into any small number of discrete categories when the interindividual variability is continuous.** Even when we recognize that variability is distributed continuously, we often feel compelled to draw “bright lines” that dichotomize the population in order to respond in a workable fashion to the infinite gradations. This practice can, depending on the shape of the function relating susceptibility to risk on either side of the bright line, cause us to focus too little or too much attention on “hypersusceptibles” as defined by a yes/no oversimplification. Fattis (2004) offers an example involving birth weight and infant mortality: he argues that by choosing a 2500-gram cutoff to define low birth weight infants meriting special attention, we fail to intervene to protect babies weighing slightly more than that cutoff, whose individual risks are indeed lower but who collectively account for roughly one-third of all deaths before age one. Similarly, Grodsky (2005) seems to agree with the suggestion Omenn (1942) made over two decades ago: that a bright line of relative risk should govern whether we treat individuals with different genotypes separately from the remainder of the population. If this cutoff was set as high as a 10:1 relative risk (the value Omenn regarded as preferable), people with ninefold-excess susceptibility would be treated as if they were no different from the norm. Such an outcome might reflect a rational balancing of the costs and benefits of providing differential treatment, but such balancing is foreclosed before it can begin when we treat a continuous variable as if it was dichotomous.

6. **Highlighting variability to create a scapegoat.** One additional misuse of information about interindividual variability involves a solid analysis of a relevant source of variation at the individual level, followed by a value-laden adjustment of the result. Arguably, this has taken place with respect to EPA’s attempts to estimate the exposure of “the individual most exposed to emissions” as required under the 1990 Clean Air Act Amendments and analogous efforts under other statutes. Beginning in the late 1980s, critics of “conservatism” in risk assessment reserved special derision for EPA’s methods for estimating the exposure to the “maximally exposed individual” (MEI); the apotheosis of this effort was probably the argument that only a “porch potato” could be exposed twenty-four hours per day for seventy years (Goldstein, 1989).
EPA may have believed that its procedure to estimate the MEI's exposure yielded a plausible estimate of the extreme tail of the population exposure distribution, perhaps because other assumptions embedded in the calculation, such as breathing rate or environmental concentration, were not in the tails of their respective distributions. Nevertheless, EPA eventually ratcheted back some of the assumptions in the MEI equation and developed constructs such as the "reasonable maximum exposure" and the "high-end exposure estimate," which EPA intended to represent the exposure of an individual between the 90th and 99th percentiles. Most observers, including EPA, have characterized this change as increasing the "realism" of exposure assessment, as if the new type of estimate was simply a more precise (or less fanciful) calculation of the same reference point. But if the reasonable maximum exposure and similar constructs truly represent exposures that between 1 and 10 percent of the population exceed (i.e., perhaps millions of individual U.S. residents), they mark a retreat from the attempt to protect against a true "worst-case" exposure. If so, it appears that focusing on the MEI exposure allowed critics to redefine the "worst case," rather than to confront the benefits and the costs of explicitly rejecting worst-case thinking.

See No Evil: Factors that Opacify the Veil of Ignorance

Although the preceding two sets of examples and discussion are relevant to the subset of human variability that influences risk to environmental and occupational pollutants, the breadth of examples should demonstrate the more general point: that human interindividual variability is often difficult to confront, to depict properly, and to incorporate appropriately into public policy, even when the variation is overt. The challenges only increase when additional research or data-gathering to verify the existence and breadth of variability must precede any analysis and management thereof.

The subset of human variability related to genotype, and the further subset that influences the risk of environmental or occupational disease, carries special baggage that reinforces the general reluctance to begin down the path of individualization. The special hurdles facing genetic susceptibility begin with public discourse about the very possibility of such variation, continue with attempts to quantify its magnitude or to ascribe particular levels of susceptibility to individuals, and ultimately im-
pode attempts to provide information and interventions tailored to individuals according to their susceptibility.

The reluctance to confront these issues is often so strong that the distinction between posing questions and taking irrevocable action based upon the answers has lost much of its meaning. Disquiet over the "tragic choices" we could make with (or without) new knowledge has morphed into revulsion about even seeking the information in the first place. The knowledge that would allow us to consider treating persons differently according to their risk, or to consider treating everyone in a new way because of knowledge about the spectrum of individual risks, is at once both tempting and abhorrent. In recent times, I would argue, abhorrence is carrying the day.

If we continue to see genetic information about environmental susceptibility as a "half-empty" proposition, a negative feedback loop may lock into place. Not wanting our informational capacity to increase could lead to our not advancing the science in ways that could reveal heightened opportunities for public health protection and diminished prospects for mischief. Arguably, this feedback is already occurring, as perhaps seen in Weinstein's chapter in this volume, where he reports that occupational health advocates are generally dismissive of the utility of toxicogenomics in today's regulatory environment.

The prospect of individualizing risk management via genetic information faces a one-two punch from factors one might oversimplify as squeamishness and foreboding. I do not use these terms to imply any unwarranted concern, only to connote two different phenomena that each may be wholly justified. First, our society has always tiptoed around some manifestations of variation that may be correlated with or caused by immutable characteristics such as race, sex, or age. Even to mention examples where analysis of traits shared by subgroups of the population has revealed apparent associations with other attributes—whether negative or positive—invokes both misunderstanding and passion.

To wade briefly into one such controversy, consider the January 2005 speech at the National Bureau of Economic Research by then—Harvard President Lawrence Summers about the underrepresentation of women in tenure faculty positions in science and engineering (Summers 2005). Summers suggested three possible causative factors, in this "probable order of importance": (1) the possible differential desire of men and women to do "high-powered intense work," exacerbated by the expectations of
employers that "the mind is always working on the problems that are in the job, even when the job is not taking place"; (2) the possibility that there are fewer women than men three to four standard deviations above the population mean in mathematical ability (because the variance in mathematical ability among men exceeds that of women, yielding more male mathematical prodigies and more dolts); and (3) the possible influence of the "old boy" network of gender discrimination and "like begets like" hiring practices. I need take no position on the merits or propriety of any of these arguments to observe that at least some of the vitriol that resulted in the unprecedented vote of the Harvard faculty to censure Summers, and that ultimately led to his resignation in 2006, stemmed from his remarks about differential variance in mathematical ability. One professor described those remarks as "reckless and undigested words based in half-baked sociobiological prejudices" (Harvard Magazine, 2005, p. 57). Apparently, at least as far back as Darwin's Descent of Man, scholars have posited valid reasons why men should and often do exhibit greater genetic variability than do women (Hedges and Nowell, 1995). I suspect, however, that even if Summers had suggested that women have even a higher mean mathematical aptitude but less variability (thus still fewer women at the extreme ends of human mathematical ability), many would still have concluded that one cannot summarize any characteristic of two genetically different groups (in this case, their respective standard deviations) without prejudging (in the pejorative sense) the worth of individuals within them. I suppose that to the extent the characteristic is value-neutral, and to the extent that the genetic difference is wholly uncorrelated with sex, race, or other attributes that define group membership, such observations would be less incendiary. Nevertheless, it seems clear that the very characterization of a group as defined by a genetic commonality makes any further elaboration about that group precarious.

Couple this aversion to genetic information that can open the door to invidious comparisons with a powerful new worry specific to toxicogenomics—foreboding about the deliberate misuse of individual genetic information—and the prospects for personalizing environmental and occupational risk management grow bleaker still. On the one hand, discovering you are at special risk confers potential benefits, such as the ability to protect yourself to a level greater than that enjoyed by those less susceptible or to claim that you deserve such protection in your workplace or community. However, the same discovery can be an un bottled
genie that opens you up to a panoply of new harms. In my view, the debate over the peril and promise of toxicogenomics has increasingly accentuated the negative, with some justification (see below). The very term “genetic discrimination” clearly has come to connote something very different from “discriminatory power” in the neutral or positive sense of discerning real differences and intervening appropriately to attenuate or even reward them. Fear of injury, illness, and death has increasingly stepped aside and let other fears play the trump card. Furthermore, in playing one set of rights (for example, the right to die of “natural causes” rather than from involuntary exposures to contaminants in the environment) against the other, I believe we are making little effort to find positive-sum solutions that might extract the benefits of genetic information without accepting the mischief, thereby ensuring that the competing values will remain at odds.17

My perspective on these competing rights stems primarily from having worked as a regulatory and enforcement official for the U.S. Occupational Safety and Health Administration (OSHA) and having seen firsthand the repeated failures of our rational effort to reduce risks to workers from chronic disease and acute injury to levels that we would consider acceptably low in virtually any other setting. I therefore tend to see a statement such as this one from Silvers and Stein (2002) as an ironic use of the term susceptible: “Workers with genetic vulnerabilities to materials found in the workplace... seem especially susceptible to rejection on the ground of inability to perform essential functions.” The authors clearly intend this as an example of the “dark side” of genetic information and frame “rejection” as the insult—which seems as one-sided as a summary that would say merely that “genetic screening can save the lives of workers who would otherwise face virtual death sentences from exposures that are singularly inappropriate for them,” without mentioning the loss of opportunity concomitant with “rejection” (see below).

Given that most OSHA standards (to say nothing of the vast majority of substances for which no OSHA standard exists) allow exposure levels corresponding to lifetime excess risks of $10^{-3}$ or greater (and as high as roughly $10^{-2}$ in OSHA’s most recent health standard [U.S. Occupational Safety and Health Administration, 2006]), even without accounting for genetic and other variation that may place many workers at even higher individual risk, I find it impossible to view “discrimination” that might
lower risks to highly susceptible individuals as solely a detriment to the person or persons involved. Although in environmental policy, typical risk levels (without regard to susceptibility) tend to be two or three orders of magnitude lower than those in the workplace, I still see the potential there to provide highly susceptible individuals with less onerous environmental exposures as a balance that must be struck between competing rights, rather than only as an infringement.38

It would be flippant, however, to conclude merely that without life and health, none of the other rights at issue can be enjoyed. Genetic and other information about individual risk potentially impinges on many substantive and cherished rights, raising concerns about

*Stigmatization.* Identifying persons at high risk of environmental disease because of genetic predisposition raises all of the concerns associated with the original meaning of “stigma”—an indelible mark branding someone as undesirable. This is perhaps the most inchoate but compelling of all the objections to individualized risk assessment, in part because the technology itself is so open-ended. As Silvers and Stein (2002) suggest, acquiescing to a single DNA test, even one where being identified as high-risk confers no stigma, generates personal data that could years later brand you, upon the discovery of a hitherto-unknown genetic marker, as more disadvantaged.10 The Catch-22 of many of the existing protections against genetic discrimination is that they provide remedies under the Americans with Disabilities Act, thereby making the label “disabled” a price of protection: (Silvers and Stein, 2002).

*Insurability.* Although the federal government (in the 1996 Health Insurance Portability and Accountability Act) and at least forty states have enacted legislation banning some uses of genetic information for denying insurance to otherwise-eligible applicants (Clayton, 2003), concern about insurance companies simply pricing such applicants out of the market remains high. Concern over insurability runs especially high because of an inherent Pandora’s box quality of individual genetic information—it can stigmatize or otherwise harm close relatives who were not aware of, or specifically did not consent to have revealed, insights into their own genetic makeup.

*Job loss.* As several authors in this volume have noted, there currently is no federal law prohibiting private-sector employers from
firing, or refusing to hire, an employee or applicant because of his or her susceptibility to disease. Leaders in the labor movement clearly view employer latitude to screen current and prospective workers for susceptibility as "another form of discrimination against workers" (Sprinker, 2005), rather than as a potential means of selectively reducing the existing discrimination that allows workers as a class to be exposed to risks a thousand to a million times higher than society generally tolerates in the ambient environment (Finkel, 2005). Indeed, the 2002 Supreme Court case Chevron U.S.A. Inc. v. Echazabal now explicitly allows employers to deny employment to anyone with a medical condition that arguably places him at higher risk of disease from exposure to particular substances in the workplace, on the grounds that a basic qualification for employment is the ability to do the job safely, even if the employee's susceptibility places no one else at risk (as would occur in the scenario of an airline pilot with episodic vertigo).20

- **Loss of autonomy.** This concern encompasses a variety of harms, beginning with threats to the simple human desire not to know or even to suspect one's susceptibility or disease status—what Diver and Cohen (2001) referred to as the "nocebo effect." This effect might be especially potent when the marker of susceptibility to future exposure implies a grave health problem even in the absence of further exposure.21 Even if subjects choose to receive the test results, subsequent decisions affecting their autonomy may be beyond their control. The general concern about having to adapt to the environment—keeping one's job but having to wear a respirator, for example—may be as daunting as the prospect of having to leave the environment entirely. Anecdotally, one explanation for the relatively small percentage (about one-quarter) of active OSHA inspectors who, at this writing, have availed themselves of the free blood tests for sensitization to beryllium (Young, 2005) is their concern that a positive test will result in their being assigned to a desk job or prohibited from inspecting those many facilities where beryllium dust could conceivably be present.

- **Debasement.** Some scholars (e.g., Wolf, 1995) believe that the flaws of individualized risk assessment and management do not depend on whether any harms are done to individuals; that a society that embarks on this path debases itself from the outset. Wolf decries
"the eagerness to draw genetic conclusions, the search for supposedly deviant genes, and the conviction that such genes actually deserve disadvantage," and suggests that the "deeper harm" is an attitude of "geneticism" that derives from repugnant instincts we should quell. She asserts that seeking and acting upon information about individual characteristics is offensive "even when based on accurate rather than exaggerated understanding of the role of genes."

In other words, a test with 100 percent accuracy, used to assess risk to a single individual without regard to any other characteristics defining membership in a group (race, sex, age, etc.), dehumanizes the society that would develop and use it.23

Although some of these harms are theoretical, it must be noted that specific recent uses of information about interindividual variability in susceptibility tend to validate the concern that the information will be sought and applied clumsily in practice. Cases such as the blood-testing policies of Lawrence Berkeley Laboratory, in which African American employees were reportedly tested repeatedly for sickle cell trait and syphilis (Silvers and Stein, 2002), and the Burlington Northern case referred to throughout this volume, in which the railroad for a time required workers filing claims for carpal tunnel syndrome to submit to genetic tests for a rare hereditary neuropathy, give credence to a pessimistic stance. These cases and others suggest that fears of misuse, racial discrimination masquerading as genetic counseling, and blaming the subsidiary cause (the victim) rather than the proximate cause (the environment) are grounded in a sober assessment of economic and social realities. These concerns are bolstered by examples from the more distant past of eagerness to distort genetic information (Gould, 1981). The combination of such specific fears, along with the general reluctance to court social turmoil, regardless of the case-by-case safeguards that might be built in, form a potent argument against using many types of scientific information to individualize public health interventions.

I saw these tensions play out as a member of the group that drafted Executive Order 13145 in 1998–2000 (Clinton, 2000), in developing a policy that prohibited any federal agency from firing, refusing to hire, or "classify[ing] employees in any way that would . . . adversely affect that employee's status" because of information about the employee's genotype (whether gleaned through genetic testing or by information on dis-
ease status of family members), there was essentially no discussion about the potential benefits of such information to heighen protections for individual employees or for the workforce as a whole. Several narrow exemptions were carved out near the end of the process, including a clause (e1) that allows OSHA to promulgate future health and safety standards that might allow or require some genetic testing, and a provision (e3) that allows the FBI and other agencies to perform genetic testing “to carry out identification purposes,” but the order contains no exemptions to protect the health of workers who might face intolerably high excess risks due to genetic susceptibility.

The evolution of subpart (d) of the order, which distinguishes genetic monitoring (the analysis of DNA and other macromolecules to examine mutations acquired during the course of employment) from other genetic testing, indicates how reluctant the other federal agencies in the task force were to “crack open the door” to a perceived genetic technology. When the group hammered out the definition of “genetic test,” it was pointed out that “analysis of human DNA . . . to detect disease-related genotypes or mutations” (emphasis added) subsumed genetic monitoring, even though that practice cannot reveal information about an individual’s or a family member’s genotype. By definition, the adducts or point mutations detected are not inborn or shared among relatives, but acquired. The group redefined properly safeguarded “genetic monitoring” as outside the definition of “genetic test,” but it nonetheless determined that genetic monitoring would not be permitted in the federal workforce unless the employer “receives results of the monitoring only in aggregate terms that do not disclose the identity of specific employees.” The question was left unanswered as to what beneficial use an employer could possibly make of the knowledge that some unknown employee had been overexposed to a workplace contaminant. After all, if the biomonitoring was “nongenetic” in nature, as has been ongoing for many years with respect to measurements of blood lead, urinary cadmium, and other substances in the workplace, we recognize that for the employer to know where engineering controls are inadequate (and to provide paid leave or medical treatment to normalize the excess body burden(s) of the worker(s) affected), she needs to know whose test results are abnormal, or at least the work area in which the readings are elevated. Here the concerns over possible misuse of employee-specific (and nongenotypic) information seem to have trumped the intended benefits of a regime with potential
for identifying risk-reducing or life-saving interventions—in a microcosm of the direction of the larger debate discussed in this chapter.

Individual genetic information thus repels many thoughtful observers, because it can have far-reaching negative consequences not only for the individuals who submit to testing, but also for their family members, the groups they belong to, and society as a whole. Even when those concerned about the potential for harm acknowledge the possible benefits forgone by restricting the exploration of individual susceptibility, they often damn with faint praise, as in this well-known quote of Francis Collins (Wade, 1998), the director of the Human Genome Project: “This ability to collect very large amounts of [information on] variation on individuals will be quickly upon us. It will empower people to take advantage of preventive strategies, but it could also be a nightmare of discriminatory information that could be used against people” (emphasis added).24

Rather than a tool to successfully demand an inherently safer workplace or community, individual genetic information is thus portrayed as, at best, information that might force the vigilant individual to work harder to bring about his own health benefit. Similarly, the positive mirror image of susceptibility—“resistance” or immunity—rarely receives attention, yet surely some individuals will benefit (e.g., with lower insurance premiums) from being able to show personal resistance to disease or exposures. Something must explain this overwhelming pessimism, and I offer two observations before proceeding to discuss possible regulatory and public health responses to interindividual variability in susceptibility.

First, perhaps we seek to honor the subset of genetic variation that we need to work hard to detect, as if the genetic code itself hides secrets for a purpose. My sense is that in many of the cases where critics reject differential treatment by genotype, their conclusions would differ if the susceptibility was overt or required no “peaking into the DNA.” To oversimplify, I think we would, perhaps reluctantly, agree that albinos (overt variation) should be discouraged from working as lifeguards or hemophiliacs (variation detectable via symptoms or via nongenetic analysis of blood) as meat-cutters—but somehow we conclude that choosing not to know about exposure-specific susceptibilities contained in the genetic code absolves us from confronting such “tragic choices.”

Second, perhaps pessimists correctly see individual genetic information as a lose/lose situation, for the following reason: if genetic testing
reveals that individual susceptibility is correlated with group membership, the desire to discriminate against the individual can reinforce or reawaken the desire to discriminate against the group. But what if—as seems more biologically plausible in many cases—susceptibility is unrelated to sex, race, age, or ethnicity? Then those identified as high risk become a de facto group, but may be “hung out to dry” because they don’t form a class we otherwise feel poignant about or because we are reluctant to appear to be discriminating against a new group.25 This may be the most far-reaching concern of all—that individual genetic information can reveal a need to protect, which society can then comfortably turn its back on, leaving only the stigma behind.

Assuming, therefore, that quantifying susceptibility at the individual level closes more doors than it opens, the management question remains: What, if anything, can we do when we know that variation in susceptibility creates a distribution of risks in the environment or workplace, but when we also believe that learning the identities of those most susceptible carries too high a price?

Does Identifiability Change Everything?

One response to the problem of protecting individuals according to their susceptibilities would involve lowering the cost of identifying individuals and reducing the likelihood that such identification would be necessary—in other words, to make identification a last resort when all other protections have failed and to surround it with additional safeguards. In the occupational setting, executive branch policy or federal or state legislation could allow for differential treatment of a highly susceptible worker only if several strict conditions were met. Such conditions might include (1) the diagnosis was reliable and conferred susceptibility to a specific exposure or exposures in the given workplace; (2) both relative risk (the degree to which the predisposition elevated the worker’s risk above that of the rest of the workforce) and absolute risk (the probability of harm given exposure) were substantial, so that any “victim-level controls” such as personal protective equipment or job transfer would apply only to workers at significant (and significantly elevated) risk; (3) the information gleaned from any genetic test was the property of the employee and could not be revealed without consent to any other party; and (4) differential treatment could not include firing or refusal to
hire, under any circumstances. Most important, the employer seeking to implement a “victim-level” control strategy would have to provide evidence that it was infeasible to further lower ambient exposures, either across-the-board or specific to the susceptible employee or employees. An even more limited role for treating identified hypersusceptible individuals differently might arise in environmental protection; perhaps warnings against excessive consumption of certain contaminated foodstuffs could be targeted primarily at individuals known through genetic testing to be at high relative and absolute risk.

These could be viewed as “stubborn” responses to the tension between rights to health and other rights, with the goal of forcing a balance between them, accepting that some may suffer losses of privacy and autonomy in the name of their own health. It may be possible, however, to sidestep this tension entirely, by trying to protect populations based on the susceptibilities of unidentified individuals within them. A “strong form” of such a policy might state that occupational or environmental standards “shall provide acceptably low risk (or ‘reasonable certainty of no harm’) for X,” where X could be defined at a desired level of inclusiveness. Some points along this continuum might include “the entire population including the most susceptible single individual within it,” “the population up to and including the average member of the most susceptible subgroup,” or “z percent of the population when arrayed in ascending order of susceptibility.” A “less strong” form might embrace cost-benefit balancing, but within this framework set monetary values for risk reduction that increase more than linearly as individual risk increases, so that estimates of monetized population risk would not “average away” the disproportionately high risks that susceptible individuals may bear.

Before proceeding any farther to evaluate the merits of either type of policy, one must acknowledge that the very notion that unidentifiable variability could affect public perception or public policy may be illogical or incoherent. Many scholars believe that unidentifiable variability in risk cannot matter, for several reasons. This view generally begins by acknowledging that identifiability clearly makes us much more concerned about high individual risks and those who face them. Jenni and Loewenstein (1997), for example, introduce their paper on the “identifiable victim effect” with a compelling observation made by Thomas Schelling in 1968, attesting to the much higher implicit value we tend to place on hu-
man life when we can “rescue” an identified person. It identifiability confers such a premium, it certainly is possible that the converse is true—that absent identifiability, the mere knowledge that someone’s risk is intolerably high is of no particular consequence, beyond the proportional effect of those risks on the total risk to the population. More formally, the most powerful argument in favor of the proposition that a distribution of unidentifiable individual risks can and should be reduced to the average probability of harm observes that each individual probability of being susceptible is itself a risk, and asserts that only the expectation of this “second-order probability” should affect the perception of harm.

Consider these hypothetical scenarios, involving two cities, each having one million inhabitants, and a proposal on the table to locate a new source emitting a carcinogenic air pollutant in each city. In city A, the inhabitants all have identical genotypes and environmental histories, such that their additional lifetime risk from the new chemical is $10^{-4}$. In city B, ten thousand of the inhabitants are highly (one hundredfold more) susceptible to the carcinogenic effects of the pollutant, such that their excess lifetime risk is $10^{-2}$, and the other 996,000 people in city B are immune and face no risk. In both cities, one hundred people are expected to die from their exposures ($10^6$ people in city A multiplied by a uniform risk of $10^{-4}$ equals 100; $10^4$ susceptibles in city B at a $10^{-2}$ risk also yields 100). It is thus plausible to argue that any individual, lacking any information on his individual susceptibility, should be indifferent between living in city A or city B; in either case, his expected excess risk with the new pollutant source would be $10^{-4}$, or the number of fatalities divided by the population size.

Jenni and Lowenstein (1997) support this view, using a medical hypothetical with slightly different parameters, when they say, “it probably makes no sense to treat a disease that kills 100% of the 10% of the population susceptible to it differently from one that kills 16% of the 100% of the population susceptible to it” (in both cases, the expected individual risk here is $10^{-1}$). Adherents to this view presumably allow for departures from risk-neutrality at the societal level. The decision maker concerned about a population of size $P$ may either prefer to accept a distribution of possible numbers of total fatalities whose mean is $X$ to a certainty of accepting exactly $X$ fatalities (risk proneness), or vice versa (risk aversion), but the individual citizen should be indifferent between an uncertain personal risk with mean $X + P$ and a risk known with certainty to
be exactly \( X + P \). Note that in the hypothetical above, the binomial uncertainty in the number of fatalities around \( X = 100 \) is small compared with \( X \) and is almost exactly the same in city A versus city B (see below).

If variability in risk has no significance until we can ascribe particular points on the risk distribution to particular persons, then as a practical matter we should feel comfortable censoring information about interindividual variability that only reveals the breadth of differences among people. The answer to people who express concern about equity as well as efficiency would simply be that “people can’t be treated inequitably if neither we nor they know who they are.” However, it would be imperative to admit that “the mean is not the message” and perhaps to control hazards differently based on the distribution of individual risk, if unidentifiable variability does matter, which I will argue it does.

Is “Someone” Identification Enough?

In contrast to the proposition that only identifiability refines human variability, several arguments support the view that the distribution itself may be as compelling as the identities of everyone described by it.

One argument does not even require the individuals themselves to care about unidentifiable variability. Even if information on the number of fatalities and its uncertainty is sufficient to make risk-management decisions, ignoring interindividual variability in risk can lead to errors and biases in estimating both the mean and variance of the “body count.” The assessor who focuses on the fatality distribution may not realize what Feller (1968) referred to as a “striking result”: the uncertainty in the number of events from a binomial process is at its maximum when the probability of success or failure is identical for each member of a population. The decision maker in city A above expects one hundred extra fatalities but should expect a standard deviation of almost exactly ten around this mean estimate. In city B, however, the number of fatalities is less uncertain (the standard deviation is approximately 9.95), because in this exaggerated case where the risk is completely concentrated in a subpopulation, there are fewer people who contribute anything to the “body count.” Therefore, a decision maker who is risk averse (or risk seeking) with respect to the total number of fatalities in the population may be less (or more) concerned about a hazard if it applies to a diverse population rather than to a homoge-
neous one. Put another way, ignoring variability known to be present leads to an exaggerated estimate of the actual uncertainty in the number of fatalities or cases of disease, even if the estimate is correct on average—so this may affect the decision even if one is wholly uninterested in the distribution of individual risks.

A different and probably more substantial problem concerns bias, rather than uncertainty, in estimating the expected number of fatalities. The central limit theorem dictates that the observed mean is a useful estimate of the true population mean, but in the presence of unmeasured variability, the errors in that estimate may be much larger and more complicated than elementary statistics suggests. Suppose that the risk of some event was strongly related to income, so we needed to know the average income of the population to estimate risk. In this case, the influence of outliers is strong and asymmetric, such that it is hard to get an accurate estimate of the mean income unless by chance the sample contains the correct proportion of very rich people. Too many billionaires and the estimate of the mean will be much too high; too few, and the opposite will be true. It turns out that for quantities such as income, exposure, and susceptibility, that are nonnegative and distributed approximately normally on a logarithmic scale, the error in sample estimates of the mean is itself lognormally distributed. It is more likely that the observed mean will be slightly below the population mean, but the largest absolute errors will be those less likely instances where the observed mean is much larger than the population mean (Finkel, 1990).30

Thus the paradox: even if the average is all that matters, it can be very difficult to find the true average without first exploring the entire distribution which contains it. This is a powerful argument for quantifying variability even when it is nonidentifiable, which would then put us in the position of having to ignore information on variability that we have already accounted for. We can still struggle over how to act on this knowledge, but it seems that treating diverse populations as homogeneous can do violence to efficiency (expected-value decision making) as well as to equity.

Beyond the value of quantifying unidentifiable variability in revealing important characteristics of population risk, various other arguments run counter to the view that individuals should be satisfied with an estimate of their mean risk. Do we really believe that situations where an entire population faces an identical probability of harm are perceptually and
ethically equivalent to situations where the mean probability is unchanged but the individual (albeit unidentifiable) probabilities spread farther and farther apart from each other? Is a disease that kills one in ten people at random truly equivalent to one that kills everyone of the 10 percent of the population susceptible to it?

A precondition for answering "no" to these questions is agreeing that individual risks themselves are not merely weights that moderate the ultimate outcome (death, disease, injury, etc.), but have salience of their own. Then, the one in one hundred chance that a resident of city B will face a risk of $10^{-2}$ is not the same as a certain risk of $10^{-4}$; not knowing whether one is at high risk or no risk is simply not the same as knowing for sure that one is at an intermediate level of risk. Put a different way, we often assign a "value of life" estimate to gauge the human cost of imposing an individual risk, the cost to each person being the "value of life" multiplied by the probability of fatality. If the magnitude of that cost has anything but a strictly linear relationship to the size of the risk, then the distribution of risks matters. Only if every risk of magnitude $kX$ is exactly $k$ times as adverse as every risk of $X$ can we reduce the distribution to an average without distorting the conclusion.

I suggest that the very notion of "unacceptable," "intolerable," or "significant" risk says something profound about how we think about individual risk: namely, that as individual risk levels rise, our level of concern may rise faster than dictated by strict proportionality. Guidance on cost-benefit analysis promulgated by the U.S. Office of Management and Budget (2003) cautions that any monetization based on the "value of a statistical life" must be limited to "small changes in fatality risk" (the technique has "no application . . . to very large reductions in individual risks"), reflecting the notion that the total harm done to a million people facing a $10^{-6}$ risk is qualitatively different from (smaller than) the total harm done to two people, each facing a 50 percent risk, even though one "statistical life" is lost in either case.

Of the various empirical and theoretical arguments that high individual risks matter even before we identify those who face them, the hardest two to decouple relate to fear and fairness. Evidence that people regard the possibility of being highly susceptible (and thus at higher risk) as qualitatively worse than an equivalent certainty of being at lower risk could reflect either what we fear or what we regard as unfair, or both. Either way, widely varying individual risks generally arouse more concern.
than otherwise identical risks reported as invariant. Lopes (1984) reported
that the degree of inequality in a distribution of risks, ceteris paribus, is
a powerful predictor of the aversion subjects report to the situations. More
recently, Ritov and Baron (1990) surveyed whether respondents would
vaccinate their own babies against a strain of flu that would kill ten of
every ten thousand unvaccinated babies, as they varied the hypothetical
probability of death from the side effects of the vaccine itself. On aver-
age, respondents would elect not to vaccinate if the side-effect risk ex-
cceeded approximately 5.5 per ten thousand. Ritov and Baron concluded
that the unwillingness to vaccinate, even when the net risk of doing so
was still negative, reflected "omission bias," the special aversion to the
chance that an act of commission, rather than one of omission, would re-
sult in grave harm. But when respondents were told that 1 percent of ba-
bies were highly susceptible to death from side effects of the vaccine (and
that the other 99% were immune), the mean risk of side effects at which
people were no longer willing to vaccinate dropped significantly (to 4.5
per ten thousand), even though respondents were told that no test could
identify which children were susceptible. When respondents were in-
stead told that the susceptible subgroup arose because "the vaccine in-
teracts with a certain chemical naturally produced by the body" in 1
percent of children, the maximum acceptable side-effect risk dropped
still farther (to 3.2 per ten thousand). The authors speculated that "the
perception of missing information can make people reluctant to act, even
when the information is unobtainable," which would explain why the
more detailed information about the cause of the susceptibility increased
the aversion to vaccination.

An alternative explanation, however, looks to empathy and public per-
ception of fairness. If a risk is found to be confined to 1 percent of the
population but has the same mean as reported before, then the individu-
al risk to the susceptibles has been reevaluated at one hundredfold
greater than before. Willingness to accept an evenly shared risk of 5.5 per
ten thousand, but not a 1 percent chance of a risk of 3.2 per hundred,
could suggest that we are repelled to think that some people face risks
that large (or that we are afraid that we are among the unlucky or unfor-
tunate). Or perhaps we simply tend to view as more dire those situations
where we are not "all in this together." When I first read Cranor's dis-
cussion of the "eggshell plaintiff" (Cranor, 1997), I concocted a mental
picture of susceptibility as akin to a fanciful situation where a group of
a thousand people is confronted with the choice between two "risk
rooms." In one room, a rain of gravel will fall from the ceiling at just
the right intensity to kill one of the thousand at random from a hemorrhage
in a crucial blood vessel. In the second room, a boulder will drop and
 crush one of the thousand. In examining why I perceive the second sit-
uation as more dreadful, personal fear is not really a satisfactory ex-
planation, because whether I was standing under the boulder or was unlucky
enough to be harmed by the gravel, I wouldn't know it until it was too
late—it has to be the removal of the shared risk that distinguishes the
two rooms. Ex post, someone in the second room turns out to have been
facing a crushing risk (literally) all along, while the rest will learn ex post
that they were not in fact sharing in this grisly lottery, and this seems to
make the one in a thousand chance of facing certain death worse than
the certainty of a one in a thousand risk.

Whether fear or fairness (or both) drives this concern, it would be
much more useful to be able to gauge how much we care about uniden-
tifiable victims in the tail of the risk distribution, beyond the observa-
tion that the information may affect us qualitatively. The answer to this
estimation problem will doubtless vary across respondents and across
situations, but probably the two most important determinants are in-
versely correlated with each other (for a constant population mean
risk, so we can be comparing situations with equal expected fatalities):
(1) the proportion of the population at highest risk and (2) how substan-
tially those high risks exceed the typical risks. For example, the popu-
lation risk in city A would also equate that in "city A" if the susceptible
subgroup consisted of 0.1 percent of the population at one thousandfold
excess risk, 40 percent of the population at 2.5-fold excess risk, and
so on. Qualitatively, I think it possible that at one extreme (the small-
est fraction are susceptible and their excess risk is maximal), concern
might diminish, on the grounds that those most affected are "too few,
too different" (or, regretfully, so frail that protecting them from one risk
would only allow them to succumb to one of many other maladies to
which they are also hypersusceptible). At the other extreme (very large
numbers of individuals at only slightly elevated risk), very minor
policy changes, if any, would be needed to account for the variation.
This suggests that it is in the intermediate cases where smashing our
myths of homogeneity in risk might be most profoundly unsettling to
the status quo.
One useful way to summarize and communicate the interplay between the breadth of a risk distribution and the probability that a random person will face greatly elevated risk is to explore the "concentration function," the relationship between the fraction of the population contained within different portions of the distribution and the fractional amount of the total characteristic (in this case, risk) within each portion. Perhaps the most familiar recent use of concentration functions involves various benchmarks of the fraction of total U.S. wealth (or tax burden) associated with different deciles and percentiles of the population (e.g., Johnsen [2005], featuring statistics such as "the share of the national income earned by those in this uppermost category [the top 0.1%] has more than doubled since 1980, to 7.4% in 2002").

Figure 17.1 makes use of a simple formula (Finkel, 1990) for the concentration function of a lognormal distribution and shows the fraction of the risk borne by three arbitrary subgroups in a heterogeneous population—those whose susceptibilities place them in the 90 percent of the population at lowest risk, those between the 90th and 99th percentiles, and those above the 99th—as a function of the amount of variability in the population. A vertical line drawn for any degree of variability divides the three groups according to their "mass," or the fraction of the total number of fatalities in the population that will befall each group. For example, at a logarithmic standard deviation of 2.0 (describing a situation where 10 percent of the population faces risks at least thirteen times greater than (or one-thirteenth as great as) the typical person, and 1 percent faces risks at least 705 times greater or smaller), about 39 percent of the fatalities will befall the 9 percent of people in the middle group. That is, if one hundred thousand people face a mean risk of $10^{-3}$, thirty-nine of the one hundred expected total deaths would occur among the nine thousand people belonging to this subgroup—and their average risk would in fact be thirty-nine in nine thousand, or more than four per thousand.33

As the variability increases, the concentration of expected fatalities in the uppermost 1 percent of the population naturally increases as well, but I think it is important to note that the consequences that befall the "reasonably highly susceptible" subgroup (i.e., the height of the dashed arrow in Figure 17.1 in the unshaded portion as it moves across the diagram) is rather insensitive to assumptions about overall variability. For all values of the logarithmic standard deviation between 1.0 and 2.6,
Figure 17.1. Fractional amount of a lognormal characteristic contained within three portions of a heterogeneous population.

...roughly one-third of all the risk is concentrated among those more susceptible than most but less susceptible than the "outliers." Perceptions of "how safe is safe enough?" therefore, may be enriched and changed if we consider that perhaps 30 percent of the benefits of more stringent controls (which is to say, 30 percent of the total harm done by stopping short of eliminating the risk) will accrue to those among us who are moderately, but not exceedingly susceptible. Nine percent of the affected U.S. population could comprise upward of 25 million people in the case of a ubiquitous pollutant.
Other, more pragmatic arguments exist in favor of the proposition that knowing the extent of variation in risk may trigger the same perceptual and policy changes as knowing the identities of those at different points on the distribution:

- **Perhaps without realizing it, we already let unidentifiable variability in exposure affect policy.** Although some exposure models allow the analyst to pinpoint the identity of the people whose exposures determine the stringency of a regulatory standard (e.g., sportfishermen who consume a given large amount of their catch), often the models produce an unidentifiable distribution and an unidentifiable reference point. For example, air dispersion models may specify the concentration of a pollutant at the geographic location where an emissions plume exerts its maximal effect but only in terms of the radial distance from the emission source without regard to direction. We learn in such cases how much exposure the MEI receives, but not who the MEI is. Evidently, this circumscribed knowledge is sufficient to drive risk management.\(^{15}\)

- **Perhaps without realizing it, we also have already let unidentifiable variability in economic burden affect policy.** To the extent that concern over the upper tail of the distribution of compliance costs can cause regulators to relax the stringency of a proposed standard, or abandon a regulatory attempt altogether, often the specific individuals or companies at risk are not identifiable. At OSHA, for example, economists would declare that a proposed exposure limit was infeasible if analysis showed that one or more broad industry classes might have to expend a significant fraction of profits to reduce exposures to that level. The answer to the question “for whom would this standard be economically infeasible?,” had anyone ever asked, would have been “someone.”

- **Congressional intent motivates attempts to protect identifiable (and probably unidentifiable) subpopulations.** The text of the major environmental statutes, and judicial comment on them, may shed light on what role variability, identifiable or otherwise, can or must play in regulation. Based in large part on Grodsky's (2005) excellent summary, one can conclude that most of EPA's statutes endorse—or even demand—that the agency provide an acceptable level of residual risk for people of heightened susceptibility.
Much of the fine-tuning of regulatory stringency EPA has undertaken [some sua sponte, some in response to court rulings] has involved the slight ratcheting back of standards that would have protected every conceivable member of the exposed population, suggesting that inclusiveness, rather than protection of the average person, is the default presumption [see, for example, Grodsky’s (2005) summary of the D.C. circuit court’s instructions that EPA provide a “plausible explanation” why its SO₂ standard would fail to protect all asthmatics undergoing strenuous exercise]. The language from the legislative history of the Clean Air Act of 1970 that the authors cite (protecting “a representative sample of persons comprising the sensitive group rather than a single person in such group”) allows EPA some discretion. This language, though, makes sense only in the context of epidemiology or controlled human experimentation, in which researchers could directly observe exposure levels that affect some individuals but do not cause a statistically significant elevation in the health status of the group they belong to. When EPA has to extrapolate from animal toxicology data to set a standard, there is no “statistically related sample” of humans to study directly, which puts EPA back in the position of having to account for human variability through risk assessment and having to articulate explicitly what fraction of the population its standards are designed to protect.

The more important question for legal interpretation is whether Congress has ever forbidden EPA from protecting subpopulations whose risk can be estimated but whose members cannot or will not be identified. To be sure, references to “identifiable subgroups” pervade environmental statutes and interpretations, as befits their drafting at a time when susceptibility was oversimplified as all-or-none and associated with other salient characteristics such as age and race. However, there appears to be no legal or policy barrier to constraining these references to include subgroups defined completely by their susceptibility, as in “people with α-1-antitrypsin deficiency,” or even “people whose ‘area under the curve’ (molecules of a toxic metabolite produced in a target tissue) puts them at the 95th percentile of the distribution of this quantity.” I assume that even the more specific language in the Safe Drinking Water Act Amendments of 1996 (“infants, children, . . . or other subpopula-
tions that can be identified and characterized”) is perfectly compatible with “identification” in the population sense—estimating the relative size of a subgroup defined by a particular predisposition and estimating the potency of that predisposition to increase disease risk—rather than in the sense of “identifying” the specific individuals at heightened risk.\textsuperscript{36}

- At least one well-known federal regulation is already based on the risk to the unidentifiable individuals at the “reasonably high” end of the distribution of biological variation in susceptibility. OSHA’s regulation of methylene chloride (U.S. Occupational Safety and Health Administration, 1997) demonstrates that quantitative information about interindividual variability in susceptibility can be amassed, analyzed, and used to set a risk-based exposure limit. In the proposed version of the methylene chloride standard, OSHA had generated a point estimate of excess cancer risk to exposed workers using positive mouse tumor data coupled with OSHA’s standard assumption that doses could be extrapolated from mice to humans via the relative body weights of the two species. In revising the risk assessment during 1995-97, OSHA decided that various teams of researchers had adequately demonstrated that a physiologically based pharmacokinetic model of methylene chloride metabolism in mice and humans better explained the interspecies differences.

Data existed, however, pointing to significant interindividual variability in enzyme kinetics and other biochemical parameters within the human population. OSHA staff and contractors used a Bayesian method to integrate published information about the uncertainty in approximately thirty pharmacokinetic parameters in mice and humans with individual data on human variability, taking special care to estimate and account for correlations between relevant pairs of parameters.

Although my OSHA colleagues and I were unable, due to limitations in the data, to completely decouple uncertainty and variability (and did not attempt to quantify uncertainty or variability in the function relating tumor response to delivered dose), we did develop a probability density function for delivered dose that described the excess risk at a given workplace concentration of methylene chloride for individuals with different pharmacokinetic
behavior (U.S. Occupational Safety and Health Administration, 1997). We asserted that it was a reasonable interpretation of the 1980 Benzene decision to consider the risk to the (unidentifiable) individual at the 95th percentile of this distribution. The risk at this reference point was approximately tenfold higher than the risk at the median of the new distribution, and approximately threefold higher than the mean risk, because the distribution was right-skewed.37

The stakes associated with the OSHA methylene chloride standard may not have been as high as in many of EPA's rule makings. However, the fact that this part of the overall controversial analysis was not challenged during congressional oversight hearings or in ultimately successful negotiations among government, industry, and labor to avoid litigation over the rule, suggests that protecting unidentifiable individuals at the tail of a risk distribution can win acceptance as sound science and as sound science policy.

Conclusion: Toward Embedded Precaution in Risk-Based Policy

Three observations have led me to identify an impasse and to suggest a particular solution to it:

1. Individual risks do matter, and reducing risks to acceptably low levels for as many individuals as possible, consistent with feasibility or other cost-benefit thinking, is the core mission of environmental and occupational health;
2. Knowing the identities of those at highest risk would ensure that we would try to protect them assiduously, but knowing merely that they exist imposes similar, if not identical, ethical duties; and
3. Even if identifiability would cement in place policies that otherwise might be "hard sells," we cannot wait for identification that may never come, because it is held back by legitimate concerns over the use and misuse of individual information related to susceptibility to disease.

In the light of these factors, I prefer a "strong form" (see above) of a science-policy response to unidentifiable variability: we should learn as much as we can about the population parameters of interindividual vari-
ation, and ratchet up our environmental/occupational health goals to acknowledge and protect those—identifiable or not—who would otherwise be left behind. The foundation of such a policy would be the assertion that above some individual probability of grave harm (Congress' various instructions to EPA setting this "bright line" at $10^{-6}$ may strain the limits of feasibility, while OSHA's clinging to the Supreme Court's uppermost benchmark of $10^{-3}$ arguably errs in the other direction), involuntary risks are presumptively unacceptable if they can be reduced further. To the extent we can do so for all citizens, regardless of their susceptibility to the effects of a particular substance or stressor, we will recognize, as Cranor suggests (this volume), that "their claim to protection is not non-existent or extinguished" because of their genetic makeup or prior exposure history. I therefore endorse the suggestion of Hathis and Anderson (1999) that agencies should try to extend the mantle of protection to a relatively high proportion of the entire population, defined either by choosing a reference point such as the 99th percentile of a continuous distribution of sensitivities or a related reference point such as "the average member of the most susceptible significant subpopulation" (Cranor, this volume).

The raw material to drive such policies must come from the examination and testing of individual human beings, coupled with sophisticated biostatistical techniques (Bois, 2001) to separate uncertainty in the measurement of biological parameters from the variability revealed in multiple measurements. Therefore, Congress should consider, as a higher priority than the recent contentious and long-delayed proposals to regulate genetic information in employment, funding an array of such studies, after establishing a set of strict protections to safeguard the rights of the smaller number of subjects needed to amass information about population variability. For example, genetic and biochemical information should be gleaned from volunteers who agree to be studied for this purpose, it should be encrypted, and it should be divulged only to the individual involved (who assents to receive the information) and to no other party for any reason, in any form that can be traced back to the specific individual tested.

With each new finding of a genetic or other source of human variation that is likely to affect susceptibility to environmental disease, we need to move quickly to gauge the prevalence of this factor in the human population and the extent to which it heightens susceptibility. The
regulatory agencies need to keep abreast of this information—if any new micromanagement of risk assessment needs to emanate from the Office of Management and Budget or Congress, it is here—for the critical purpose of determining whether existing controls adequately reduce risks to an acceptably large fraction of the population.

I can also easily imagine two “less strong” forms of this policy that might be useful. First, if it is crucial to gauge the monetary benefit of providing acceptable risk to a large fraction of the population, we could accommodate susceptible individuals into a cost-benefit framework, simply by abandoning the fiction that benefit is strictly proportional to the reduction in the average probability of harm. The benefits of eliminating emissions in city B exceed those in city A by an amount we could estimate in light of econometric research, because the benefit to each person freed from a $10^{-2}$ risk exceeds the benefit to one hundred people freed from a $10^{-4}$ risk. Second, in many situations we might be able to provide equivalent protection at reduced costs by controlling risks at the “victim level,” but in a way very different from the discriminatory methods (e.g., job reassignment) that critics of genetic screening focus on. Rather than identifying hypersusceptible individuals and, for example, providing them bottled water as an alternative to cleaning the groundwater they drink from, we could encourage such “victim-level” controls across-the-board, even if they are necessary for only a small fraction of those who receive them, with the consent of the affected population. This would merely substitute one infringement for another, of course, but here we would be inconveniencing (or worse, I acknowledge) a majority in order not to identify the minority.

A policy goal of protecting unidentifiable individuals may ultimately be an accommodation to the concerns about discrimination, without surrendering to Kierkegaard’s concern (see fn 23) that we “heap [people] in a mass and defraud them.” Such an approach may be the wave of the future, not only in the environmental and occupational health arena. I see a close analogy between the topic of this volume and the observations Cook (2005) made about medical care and insurance: “It is precisely this danger [genetic discrimination], however, that may lead to a great breakthrough: the inevitable movement to universal health care…. Only [thus] will we be able to pool risk for the entire country and share what nature has dealt us; only then will there be no motivation for anyone or any organization to ferret out an individual’s confidential, genetic makeup.”
Sullivan (2000) said the same thing more pointedly: "In the long run, only the government will be dumb enough or enlightened enough to mandate a national insurance pool that works." Substitute "mandate a system of environmental and occupational health standards that protects people without punishing them for their genes," and you have one prediction about the direction the toxicogenomics revolution may take us.

In either health care or environmental protection, the most serious looming threat to such an evolution seems to be the possible backlash from citizens who realize, as they should, that protecting the tail of risk distributions is tantamount to "overprotecting" the majority. Robert and Smith (2004, p. 511) allude to this, in the context of victim-level controls that remove susceptibles from exposure, by stating "we may witness pressure to increase the tolerable level of toxins in the environment," but that same concern has always applied to across-the-board controls that might increase economic costs to many in the name of protecting the few.

In trying to predict whether such a backlash will occur, I am partly comforted by the track record of overestimation of regulatory costs (Goodstein and Hodges, 1997), which suggests that the price of "overprotection" may not be nearly as high as agencies and the regulated community initially portray. But fairness and fear, especially if they reinforce each other, may be even more important than cost in determining the overall tenor of environmental protection in the age of toxicogenomics. On fairness, echoes of support for protecting unidentifiable individuals can be found in an ancient source: Maimonides’s "eight degrees of charity" (Sacks, 2005). Other than giving someone a job so they no longer need charity (the highest degree), the highest form of charity, according to Maimonides, occurs when neither the giver nor the recipient know each other’s identities. The practical downsides of identifying those most in need of protection thus encourage us to embrace a system in which we all, depending on the hazard, can participate in this rarefied form of altruism.

In our field of endeavor, with the new knowledge that "genetics loads the gun" for each one of us as we live in a sea of environmental stressors, the line between fairness and self-interest has begun to blur more and more. Many observers of risk assessment and management, notably Justice Breyer (1993), have interpreted the following passage of John Donne (1624/1839) as an appeal to altruism: "Any man’s death diminishes me, because I am involved in mankind; and therefore never send
to know for whom the bell tolls; it tolls for thee." But in light of genetic and environmental variability, I have always read this passage as an appeal to self-interest as well and have found that the original title of Donne’s “Meditation 17,” from which this passage is drawn, supports that reading. That title, “Nunc Lento Sonitu Dicant, Morieris,” means “now this bell, tolling for another, says to me ‘thou must die.’” So “the bell tolls for thee” not only because humanity is morally interconnected, but also because our risks and our vulnerabilities are shared. Later in this meditation, Donne provides the ultimate argument for the “strong form” of reducing unidentifiable risks: “by this consideration of another’s danger, I take mine own into contemplation, and so secure my self.”

ACKNOWLEDGMENTS


NOTES

1. The examples used in this chapter reflect many different types of human interindividual variability: differences in exposure, susceptibility, economic burden, preferences, and so on. Most of the prescriptive discussion will deal with variability in susceptibility, defined here as a predisposition—due to genotype or other factors such as health status, nutrition, concomitant exposures, and the like—to disease caused by specific exposures in the workplace or the general environment. Occasionally, I will further distinguish “genetic susceptibility” from susceptibility governed by those other factors.

2. All of the analysis and discussion of this issue raised in this article occurred before the events of September 2001, which have forever changed the way we think about “groundlings,” so this example must be viewed wholly apart from that tragedy. The cited articles did not specify “unintentional” airplane crashes, although the data therein still apply to fatalities caused by factors other than terrorist acts.

3. Although at one point Goldstein et al. say that the risk was “shown to be in the range of $10^{-5}$ to $10^{-6}$ lifetime,” in context this clearly seems to be an attempt to calibrate the risk with reference to the subset of risks at this rough order of magnitude, not a statement about the uncertainty or variability in the risk. The authors refer to this range as “a range in which our society seems to have achieved a consensus that governmental action to protect public health is appro-
priate in environmental matters, yet well below the range of usual everyday risks to which we react on a personal level."

4. Although the "groundlings" example involves exposure to safety hazards, the literature on exposure to environmental hazards is replete with similar population-based thinking. For example, Mossman (1997) concludes that "radiosensitivity is a minor host factor in carcinogenesis," on the grounds that "cigarette smoking and diet each account for 15-1a times as many cancer deaths as ionizing radiation." But to an individual person who does not smoke but who is exposed to appreciable levels of radiation in the workplace or the home, the national "body count" matters not at all—from his vantage point, radiation may be far more important.

5. In fact, earlier versions of the guidelines attempted to justify the decision not to adjust individual or population risk estimates to account for human variation in susceptibility by citing studies (Allen, Crump, and Shipp, 1988; Goodman and Wilson, 1991; Hoel and Portier, 1994) showing that linear extrapolation using animal data tended on average neither to overestimate nor underestimate expected cancer death rates for the same substances as verified by epidemiologic data. A process that accurately estimates the population risk cannot, of course, overestimate risk for all or nearly all members of that population; it must underestimate risk to many. EPA may have realized that the conclusion and the explanation contradict each other: in the final version of the guidelines, only the conclusion remains, as if the explanation was a Cheshire cat.

6. We transferred two blastocysts and were blessed with one healthy daughter.

7. Aversion to one possible outcome over another is, of course, another factor that varies dramatically across individuals. Guidelines that assume a single set of preferences for every patient can impel "irrational" decisions. For example, the recommendation that pregnant women undergo amniocentesis when their risk of carrying a fetus with Down syndrome exceeds the risk of a procedure-induced miscarriage implicitly assumes equal aversion to each negative outcome (Harris et al., 2001).

8. Lack of interest in the distribution of costs may also reveal a fundamental illogic in cost-benefit policy. In theory, actions that benefit "winners" in total more than they harm "losers" are "potential Pareto-optimal" because the winners can compensate the losers such that everyone's benefit increases (or does not decrease). Many concerns have been raised about whether the expected compensation ever occurs in practice (Ackerman and Heinzerling, 2004)—but without taking an interest in the distribution of costs (i.e., identifying those who should provide compensation and those who ought to receive it), voluntary or mandatory compensation can't even be contemplated in theory.

9. In general, the more ubiquitous the product or process whose externalities (pollution) are controlled, the more the costs of control tend to be shared across the entire population. More citizens participate in spreading out the costs of reducing pollution from gasoline combustion than from golf club manufacture, so
the implicit assumption that costs are shared universally and equally is more dubious in the latter case.

10. Perhaps ironically, economists working within the risk analysis system have kept the pressure on for decisions informed by quantitative risk determinations rather than by nonrisk measures such as the reference dose or the “margin of exposure” (Workshop, 2006).

11. Note that this phraseology may also encompass subsets of homosexual men whose risks are even lower than that of the general population, although perhaps it is significant that this category is not specifically acknowledged.

12. A similar example from the environmental and occupational arena can be seen in the debate over whether EPA should be required to use “real data” on the number of years the average resident lives in his home before moving to another location (Hamilton, Vincit, and Dockins, 1997), rather than assuming exposure for a seventy-year life span. The analogous issue at OSHA contrasts the standard assumption of a forty-five-year working lifetime with data indicating that the average worker changes jobs five or more times during a career (Barmaster, 2000). In either case, advocates for a data-driven approach often seem oblivious to the fact that changing jobs (or homes) is not tantamount to changing exposures. If we allow every point source of chemical X to emit seven times more of it because people live near the source for only one-seventh of the lifetime, on average, then only those fortunate enough to move to pristine areas in every subsequent move will not see their risks rise by up to the factor of seven (see NRC, 1994, pp. 217-18, for a recommendation that EPA not succumb to this fallacy). The same problem holds in the workplace, where the relevant variable is years of exposure, not years of employment at any particular establishment or job title.

13. Virginia's algorithm sums nine subscores to generate a total score of between 3 and 76. Factors leading to a higher total score include: larceny as the precipitating offense; offender being male, younger than thirty, not regularly employed, or unemployed; prior arrest within the previous eighteen months; and number of prior incarcerations (especially if both as a juvenile and as an adult). As a result, a twenty-nine-year-old unmarried, unemployed male would start with thirty-six points (above the cutoff), even if he had no prior criminal record.

14. For exposures throughout the life span, the net effect of this three-tiered potency system is to multiply risk by a factor approximately 60 percent above its value under the assumption that susceptibility is not heightened in younger persons. The tenfold factor applies for 2 × 70 of the life span, and the threefold factor for an additional 13 × 70 of the life span: 10(2 × 70) + 3(13 × 70) + 1(55 × 70) = 114 × 70, or 1,653. As the main text suggests, however, a factor of 1.6 may be dwarfed by the actual variation among children or among adults.

15. “Typical” in this context refers to individuals who don’t display profound (and often obvious) characteristics, such as Down syndrome, that can inflate susceptibility to cancer even more markedly than can combinations of less profound predispositions.
16. Certainly twenty-four hours is at the 100th percentile of the distribution of hours per day of possible exposure. In many other cases, however, parameters criticized for being "outside the distribution" have (or already had) been found to underestimate the extreme value in the population. Many of EPA's early assumptions about food consumption patterns were later found to underestimate consumption among defined subpopulations. Similarly, OSHA risk assessments continue to use forty-five years as the maximum time that any employee can possibly be exposed to a particular hazardous substance in the workplace, even though the agency routinely receives public comment from workers and their families attesting to longer exposure periods (with the same employer, or in a series of jobs with identical exposures).

17. In a recent presentation (Finkel, 2007), I argue that the brief history of policy analysis of a similar issue—the emphasis on the possible risk-increasing side effects of interventions to reduce environmental risk—has also been marked by the pitting of one set of problems against another with a bias toward inaction. Inevitable risk-risk trade-offs may strongly argue against trying to control a primary risk, but the concern over side effects could instead motivate the search for actions that reduce both the primary and offsetting risks.

18. If, as various analyses suggest, environmental health risk assessments may underestimate risk to highly susceptible individuals by a factor of fifty- or one hundredfold, risk-based standards that reduce average risks to $10^{-3}$ may mask individual risks in the $10^{-3}$ range. At that high level, weighing the probability of infringement against the possibility of serious physical harm may cease to be a speculative trade-off.

19. This could still be a danger even if the original sample was destroyed, as the new marker could be so highly correlated with the original one that the former result implies the latter.

20. In this case, Mr. Schonental had a chronic condition (hepatitis C), and his diagnosis did not rely on any genetic information.

21. A subcurrent in the literature on the harms of testing for genetic (or other) predisposition to disease argues that the uncertainty in the information gleaned from such tests makes them particularly harmful. This argument seems to me analogous to a fallacy in the risk assessment arena: that imprecise statements about uncertainty can possibly be worse than no acknowledgment of uncertainty at all.

22. For example, the test for susceptibility to malignant hyperthermia, a life-threatening reaction to many common surgical anesthetics, also appears to reveal that the susceptible person may well develop one of several other rare neuromuscular disorders, whether or not he is ever exposed to the anesthetics (Mathews and Moore, 2004).

23. This current of thought strikes me as a very pessimistic outlook on human interindividual differences. In addition to the potential for providing life-saving information, individual genetic testing could conceivably elevate rather than debase society. Kierkegaard (1849/1941, p. 160), in The Sickness unto Death, sug-
gested that failing to admit differences among us may debase us in a different way: "And, oh, this misery, that...as for the masses of men, that people employ them about everything else, utilize them to generate the power for the theater of life, but never remind them of their blessedness; that they heap them in a mass and defraud them, instead of splitting them apart so that they might gain the highest thing."

24. When one of the prime developers of a new technology is in the vanguard of those cautioning against its use, technological optimists need to take particular notice.

25. Consider, for example, the collection of men and women of all ages, races, and ethnicities who have an otherwise silent point mutation that may make them particularly prone to developing chronic beryllium disease (McCandless et al., 2004).

26. "Let a six-year-old girl with brown hair need thousands of dollars for an operation that will prolong her life until Christmas, and the post office will be swamped with nickels and dimes to save her. But let it be reported that without a sales tax the hospital facilities of Massachusetts will deteriorate and cause a barely perceptible increase in preventable deaths—not many will drop a tear or reach for their checkbooks" (Schelling, 1988).

27. Later in this chapter, I will refer to this hypothetical again in the context of continuous distributions of susceptibility. If the one million inhabitants of city C had susceptibilities distributed lognormally, with a logarithmic standard deviation of 2 and a median risk of $1.35 \times 10^{-5}$, the average individual risk ($10^{-4}$) and the expected number of fatalities (one hundred) would also align with those in the other two cities.

28. I deliberately phrased this summary to emphasize the "tree falling in a forest with no one to hear it" quality of this argument. Not being able to identify the people facing inequities doesn't erase the disparate result—it simply obscures it. However, because so much of risk management is wrapped up in how we perceive reality, what we discern is also important. I believe that the "inequality we can't pinpoint can matter" argument is plausible but perhaps somewhat self-fulfilling.

29. For example, flipping one thousand fair coins should yield approximately five hundred heads, but the standard deviation of that number will be approximately $10$ (the square root of $npq$, where $n = 1000$ and $p$ (probability of heads on each flip) and $q$ (probability of tails) both equal $1/2$—so it is very possible that more than 516 or fewer than 484 heads will appear. But if you have one thousand unfair coins whose mean probability of heads equals $1/2$ (distributed equally among one hundred different probabilities of heads, ranging from .03 to .99), the standard deviation of the expected number of heads (which is still five hundred) will be only about 13 rather than 16—showing the (perhaps) counterintuitive property that the more different the members of the population are, the less variable the expected number of events in the population will be.

30. This property is consistent with the central limit theorem, as lognormal distributions asymptotically become normal in shape as their standard deviation
decreases (which is what happens as the size of the sample whose mean is estimated increases).

31. A strong contrary view can be found in Adler (2005), who argues that we should ignore all estimates of individual risk, because until risk finds its fruition in harm, no harm has been done. He constructs an example in which "P" is asleep in his bed while an intruder creeps to his bedside, spins the chamber of a revolver loaded with one bullet, and pulls the trigger to find that an empty chamber was in position. According to Adler, "no pair of possible worlds differing merely (in P's level of frequentist risk) are different for P's welfare (and obviously, then, they are not for anyone else's either)." By this logic, it seems to me, we can conclude that a risk was unacceptable all along (at least in the case of fatality risks) only after the victim has died.

32. In general (unless all relevant relationships are linear), the expected value of a function is not equal to the function evaluated at its expected value. So for the function relating risk level to cost (or to "disutility"), if we believe that a risk of one in ten is more than one hundred times worse than a risk of one in one thousand, then a gamble between those two risks is worse than a certain prospect that the risk is at its expected value.

33. Note that the average risk in this subgroup is only four times higher than the overall mean, because for this distribution the mean is itself roughly seven times higher than the median. Also, all of the descriptions in this paragraph could be recast in terms of income or wealth, as in "39 percent of the income in this population accrues to the 9 percent of persons whose income is between 13 and 105 times that of the typical person."

34. To provide some reference points, the logarithmic standard deviation of personal income in the United States is approximately 0.7 (and for personal net worth it is about 1.0); the logarithmic standard deviation of susceptibility in the methylene chloride example discussed below is approximately 1.5, and independent estimates by Hatfield and Barlow (1996) and Finkel (1995a) suggest that the logarithmic standard deviation of susceptibility to a typical carcinogenic substance may be on the order of 2.0. For all of these cases, therefore, roughly 25 to 30 percent of the characteristic (income, wealth, risk) is concentrated within the 9 percent of the population in the lower nine-tenths of the topmost decile.

35. Another way to look at this is to recognize that some important determinants of susceptibility actually affect the extent of exposure to individuals. As the National Academy of Sciences / National Research Council Committee on Risk Assessment of Hazardous Air Pollutants pointed out [National Research Council, 1994, p. 216], "the individual most exposed to emissions" may in fact be someone breathing less of a particular pollutant than the MEI, but whose metabolizing enzymes actually deliver more of the pollutant to target cells.

36. I believe OSHA's congressional mandate is even clearer than is EPA's in this regard, with a highly prescriptive statute that sets a goal of considering 100 percent of the affected population. To those who would argue that OSHA may not be permitted to set standards to protect hypersusceptible employees (Bergson, Campbell,
and Boszof, 2002), I asked (Fintel, 2003), in effect, “What part of ‘no employee will suffer material impairment of health or functional capacity even if such employee has regular exposure to the hazard dealt with by such standard for the period of his working life’ (emphasis added) don’t you understand?” Read in the context of the Benzene decision’s definition of “significant risk” as the probability of harm to an individual (not as the number of cases in the population, which the justices certainly could have chosen as the guide to distinguishing between significant and insignificant threats), it seems obvious that OSHA can reject a standard that eliminates significant risk for every employee if it would be impossible, but not because of any hint that OSHA should be satisfied with protecting employees of average susceptibility.

37. The permissible exposure limit (PEL) of 25 parts per million (ppm) proposed in 1991 and the final PEL promulgated were identical. In part, this occurred because OSHA’s economists had, inexplicably, not analyzed the costs and feasibility of a lower PEL in 1991, even though the risk at 25 ppm estimated in the proposal was 2.3 per 1,000 using body-weight extrapolation. Although the pharmacokinetic analysis indicated a slightly higher risk (3.6 per 1,000) at the 95th percentile reference point we chose, it would have prolonged this long-delayed rule even further to re-propose the standard with an economic analysis for a 10 ppm limit, the level at which estimated excess risk would have approached the 1 per 1,000 ceiling suggested by the Benzene court. Any comparison between the old and new risk estimates would be precarious, because both the method and the output (a distribution versus a point estimate of unknown “conservatism”) changed so dramatically between the proposed and final versions.

38. My own preference is to concentrate society’s effort on risk reduction rather than on ex post compensation, although I see various merits in Capor’s “hybrid proposal.” I believe his proposal may have an illogical aspect, though, in that a susceptible person damaged by exposure deserves compensation in his scheme, whereas an average person damaged by exposure does not. If a priori risks are made acceptably low for both people, it seems odd to distinguish between someone unlucky enough to develop disease despite low risk from someone unlucky enough to do so at a somewhat higher underlying probability.

39. One statistical wrinkle needs to be anticipated here: for some small data sets exhibiting very large interindividual variability, very high percentiles of the fitted distribution may exceed the maximum individual measured value. This is not a defect in the estimation procedure, but one of its strengths.

REFERENCES


