Working Paper Series

Working Paper #13

Balancing Uncertain Risks and Benefits in Human Subjects Research

Richard Barke

October, 2005
Revised June, 2007

School of Public Policy
Georgia Institute of Technology
D. M. Smith Building
Room 107
685 Cherry Street
Atlanta, GA 30332 - 0345

Balancing Uncertain Risks and Benefits in Human Subjects Research

Richard Barke
School of Public Policy
Georgia Institute of Technology
Atlanta, GA 30332-0345 USA

Ph. 404-894-8282
Fax 404-385-0504
Email: barke@gatech.edu

October, 2005
Revised June, 2007

Balancing Uncertain Risks and Benefits in Human Subjects Research

Composed of a variety of scientific and technical experts and lay members, thousands of research ethics committees -- Institutional Review Boards (IRBs) in the US -- must identify and assess the potential risks to human research subjects, and balance those risks against the potential benefits of the research. These assessments are laden with uncertainty, however. IRBs handle risk and its uncertainty by adopting a version of the precautionary principle, which is largely suggested by the Belmont Report and the Common Rule. To assess scientific merit, IRBs employ a tacit "sanguinity principle," which treats uncertainty as inevitable, even desirable, in scientific progress. In balancing human subjects risks and scientific benefits, IRBs use uncertainty as a boundary-ordering device that allows the mediation of science and ethics aspects of their decisions. One effect is the entangling of methodological and ethical review, which commissions in the US and UK have suggested should be more clearly separated, but decisions by research ethics committees depend in part on the negotiating space created by balanced approaches to uncertainty.

KEYWORDS *uncertainty; risks; human subjects research; ethics; benefits of science*

All policies require tradeoffs. Particularly in matters related to safety, environmental protection, and health, policy tradeoffs usually require some comparison of benefits and risks, and these usually entail both scientific knowledge and ethical issues. But there is no widely accepted theoretical or practical framework for reconciling questions about the interaction of science and values. Studies of science, policy, and society often reduce to questions about the abilities of scientific experts and non-expert citizens to navigate at the boundaries between science and the public, between fact and value, between knowledge and uncertainty, and between scientific and social processes of judgment.

In decision situations where an explicit co-mingling of science and ethics is required, we have an opportunity to explore the conceptual devices that permit accommodation between these realms. In this article I examine an organization that polices human subjects research. Nearly every nation is a signatory to the World Medical Association's Declaration of Helsinki ("Ethical Principals for Medical Research Involving Human Subjects"),[i] which generally has been extended to cover social and behavioral research as well. The US version of the Helsinki-(and subsequently congressionally-) mandated research ethics committee is the Institutional Review Board (IRB). The IRB is a peculiar quasi-governmental entity in that it has state-delegated power to alter or halt a research project,[ii] applying a set of broad ethical guidelines and regulatory mandates that are interpreted by a mixed unelected group of scientists and nonscientists in an effort to assess risks to human subjects and relate them to potential scientific benefit. Each year thousands of research ethics committees and IRBs render judgments regarding hundreds of thousands of research proposals. Many research practitioners are unhappy with the current system, in large part because of the often opaque bureaucratic demands imposed on researchers, the variability with which IRBs interpret their instructions, and concerns about "mission creep" as IRBs allegedly extend their review of research protocols into issues tangential or irrelevant to human subjects risks (c.f., Wald, 2004; Gunsalus et al., 2006).

Many of these complaints about IRBs relate to ambiguities. At the core of IRB decisions are uncertainties about what is expected by law, about how to interpret researchers' protocols, about how human subjects are likely to respond (medically or behaviorally) to proposed interventions, and about the likely outcomes of the research. My focus here will be on the treatment of uncertainty in assessing the relationship between risks to human subjects and benefits to science. Uncertainty in the IRB process is fundamental, and understanding its treatment should shed light not only on IRB decision making but also on other issues surrounding uncertain risks and benefits that arise at the intersection of science and ethics.

There are two somewhat distinct but also overlapping domains of human subjects research in which uncertainty plays somewhat different roles. In medical research benefits could in principle accrue to both society and to the individual research subject who should gain from treatment in a clinical trial. In non-clinical scientific research (e.g., social and behavioral research) it is unlikely that a human subject would directly experience any individual

benefit from participation. In this article I focus on uncertainty about benefits from the latter type of research. Medical ethicists often use several principles specific to clinical trials to compare and evaluate the benefits and risks of therapeutic research. As I will describe later, these concepts ("the uncertainty principle" and "equipoise") can be related to my analysis, but my focus is on the evaluation of scientific benefit rather than medical benefit.[iii]

Scientific knowledge is formed in an environment of inherent and accepted uncertainty. Many – but far from all -- uncertainties that accompany the production and interpretation of scientific knowledge are codified in concepts such as "95 percent confidence intervals," "margin of error," and "error tolerance." On the other hand, uncertainties about ethics and risk – in this case, about the welfare of research subjects -- are not solely within the domain of science and cannot be approached in the same way. IRBs must judge whether a risk adheres to principles of reasonableness and fairness, then weigh it against the speculated merit of the proposed research: in short, is the (uncertain) risk justified by the (uncertain) potential advances in knowledge? I will argue that IRBs handle uncertainty about risk by implicitly adopting a variant of the precautionary principle. How they handle uncertainty about science - specifically, "scientific merit" -- is more problematic, and I suggest that an uncertainty-tolerant "sanguinity principle" shapes the IRB members' interpretations of scientific outcomes. How uncertainties about scientific benefits and human subjects risks are treated by IRBs has a significant impact on their decisions, and sheds light on how scientists and non-experts engage with questions at the boundary between science and values.

INSTITUTIONAL REVIEW BOARDS

In the US, regulations that govern human subjects research (and stipulate additional protections for fetuses, pregnant women, prisoners, children, and in vitro fertilization) were endorsed in 1991 as a "Common Rule" by a wide range of federal agencies conducting or sponsoring human subjects research. The Common Rule requires research institutions that receive any federal funding from sixteen agencies and departments (including those that fund nearly all biomedical, behavioral, and social science research in the US) to establish Institutional Review Boards to assess research protocols involving human subjects. There are more than 6,000 IRBs in the US, in hospitals, academic centers, government agencies, and as independent bodies. A research institution that violates the Common Rule, even if no severe harm to human subjects results, may face severe sanctions such as withheld approval of new studies or termination of current studies, disqualification of investigators, or even suspension of all of an institution's federally-supported research. Other penalties could occur in civil courts, in lost political support for research, and in public opinion. Thus, an important function of the IRB is to protect research institutions from themselves.[iv]

There is a surprising lack of systematic empirical data about IRB structures and behaviors, with few surveys or other studies of IRBs other than performance evaluations.[v] Thus, the total scope of the impact of IRBs is difficult to estimate, but a 1998 report found that a sample of 491 American IRBs had conducted approximately 284,000 reviews in one year (NIH, 1998). At some IRBs, members read dozens of pages of detailed text for many dozens of protocols each month, then judge compliance with ethical and regulatory requirements (Keiger and De Pasquale, 2002). Few research proposals are rejected outright by IRBs, some of which approve 90 percent of proposals on an exempt or expedited basis and discuss fewer than 5 percent in detail, but one study found that fewer than twenty percent are approved as submitted; thirty-seven percent of IRBs had used their authority to suspend or terminate approved research (NIH, 1998, V-10).

Ambiguities and discretionary interpretations pervade the IRB process. In the IRB system a "human subject" is "a living individual about whom an investigator conducting research obtains data through intervention or interaction with the individual, or obtains identifiable private information." But a person queried by a sociologist is a "human subject" under the federal definition while the same person asked the same question by a journalist or a historian might not be. Similarly, the Common Rule stipulates that "research" is "a systematic investigation *designed to* develop or contribute to generalizable knowledge" (emphasis added). A sociologist who chances upon and observes interesting behavior at a social gathering conference is not conducting research, but if she attends a party with a structured protocol for observing interactions, that is "research."

When deciding whether to approve human subjects research, an IRB must first determine that the "risks to subjects are minimized" and that the research procedures "do not unnecessarily expose subjects to risk," then decide whether

the risks "are reasonable in relation to anticipated benefits, and the importance of the knowledge that may reasonably be expected to result." The Common Rule continues:

> In evaluating risks and benefits, the IRB should consider only those risks and benefits that may result from the research (as distinguished from risks and benefits of therapies subjects would receive even if not participating in the research). *The IRB should not consider possible long-range effects of applying knowledge gained in the research* (for example, the possible effects of the research on public policy) as among those research risks that fall within the purview of its responsibility (45 CFR 46.111.1, 111.2, emphasis added).[vi]

While this provision eases the IRB's need to foresee all risks and benefits, it does not alleviate its challenges in assessing the long-range risk-benefit tradeoffs in a proposed research protocol. Each IRB must determine within its loose guidelines how to balance its tasks of promoting ethics (by minimizing risks to research subjects), promoting science (by emphasizing methodological quality and research benefits), or promoting institutional objectives (by minimizing procedural delays and costs).

ASSESSING RISK

In this context "risk" refers to the likelihood and magnitude of harm or injury that might result from participation in a research study. These harms may be physical, psychological, or social (such as embarrassment or loss of employment or insurability). The task of the IRB is to determine which kinds of harms might occur, their likelihood, and whether they are justifiable in light of the rights of the human subjects and the likely benefits to science. The principles that guide justifiability are derived from the Belmont Report (1979):

> - **Respect for persons**: individuals should be treated as autonomous agents who deserve appropriate information in order to make informed judgments; those with diminished autonomy are entitled to special protections
> - **Beneficence**: researchers have responsibility to "go beyond strict obligation" and do no harm to human subjects if possible, and to minimize possible harms and maximize possible benefit
> - **Justice**: research subjects should be selected according principles of fairness and equity, not convenience, social class, etc.

An important concept in IRB decision making is "minimal risk," defined in the Common Rule as "where the probability and magnitude of harm or discomfort anticipated in the proposed research are not greater, in and of themselves, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests." This definition is intended as a sorting mechanism, determining whether all members of an IRB must conduct a detailed "full" review of a research protocol, or may evaluate the proposal using an "expedited" review that is delegated to the IRB chair or another designee.

The language of the regulations, however, provides an ambiguous standard for minimal risk. For one thing, "(i)t is unclear whether this applies to those risks found in the daily lives of healthy individuals or those of individuals who belong to the group targeted by the research" (National Research Council, 2004: 53-54). Standards to guide the interpretation of the minimal risk standard have evolved -- for example, pharmacological intervention is now automatically considered as beyond minimal risk, while routine venipuncture (blood drawing) meets the definition. Many ambiguities remain in practice, and the IRB's latitude in interpreting this concept is a frequent source of frustration for researchers and IRB members.

The Uncertainty of Risk

Risks have many dimensions. Slovic (1987) found complex psychological considerations in the way individuals perceive risk, and the cultural aspects of risk were explored by Douglas and Wildavsky (1982), who discussed the effects of shared values and beliefs on how social groups perceive risk. Among these groups are the lay public and scientists; Groth (1991) argued that scientists assess risk through the language and procedures of science, focusing on the nature of the harm and its probability, as well as the number of people who might be affected, while

Sandman (1987) found that lay citizens are less concerned than experts with quantitative aspects of risk and more concerned with attributes such as voluntariness, avoidability, and familiarity. Covello's summary of research on risk found at least 47 factors that influence risk perception (1992). Medical ethicists have addressed risk in clinical trials and developed various typologies and guidelines. For example, Sackett (2001) analyzed three levels of uncertainty in clinical trials: community uncertainty (whether expert practitioners believe the medical literature supports genuine uncertainty about the efficacy of an intervention, or "clinical equipoise"), practitioner uncertainty (the judgment of an individual practitioner about the participation of a particular patient), and patient uncertainty (how a truly informed and consenting patient evaluates the purported risks and benefits of alternative treatments). IRB members generally are not aware of this vast literature, but they understand that their task involves "inference options" (NRC, 1983).[vii]

In a typical review IRB members perceive and discuss risk more concretely. Is there a chance that X will happen to a research subject? Is that chance large or small? Is the risk avoidable? How harmful would X be if it occurred? Does the informed consent agreement make the risk of X sufficiently clear to the subject, and will the subject's agreement to participate truly be informed and voluntary? Implicit in these judgments is the issue of whether X is a "reasonable" risk, which under the definition in the Common Rule means comparing the risk of X to what is "ordinarily encountered" or "routine." The "justice" criterion of the Belmont Report asks the IRBs to consider also whether the risk of X varies across subpopulations, whose background risks may vary widely. Furthermore, different approaches to evaluating risks are likely but usually not articulated by different types of IRB members, including experts. IRB members with medical experience tend to exaggerate psychosocial risks and to downplay medical risks, while social scientists are less likely to see significant risks in behavioral research (Lane, 2005). In a survey study of scientists' interpretation of uncertainty in low radiation-dosage models and in their application of knowledge and uncertainty when recommending safety standards, scientists' values and beliefs were found to play a significant and predictable role (Silva, Jenkins-Smith, and Barke, 2007).

In addition to confronting empirical uncertainties about harms and likelihoods in a research proposal, IRB members must judge the applicable meanings of ambiguous terminology in their mandate. IRBs often adopt a device that tacitly acknowledges the epistemological challenges of analyzing possible risks to human subjects yet helps them avoid paralysis from making uncertainty the core of every consideration.

The Precautionary Principle

In practice, IRBs generally have chosen to minimize the possible harm to human subjects by tacitly applying a version of the precautionary principle (PP), the essence of which is "to take action despite uncertainty" (Goldstein, 1999).[viii] The precautionary principle involves basic concepts such as:

- "promoting the cause of intrinsic natural rights" (analogous to the Belmont Report's "justice" and "respect for persons"),
- "proportionality of response to show that the selected degree of restraint is not unduly costly" (as in the Common Rule's mandate to balance risks to subjects against benefits to science and society),
- "duty of care, or onus of proof on those who propose change"(reflected in IRB procedures requiring researchers to justify any risks to human subjects), and
- "preventative anticipation: a willingness to take action in advance of scientific proof of evidence of the need for the proposed action," which embodies the logic of precaution (O'Riordan and Cameron (1994).

Montague (1998) expanded the approach of the PP to include decision making processes that are "open, informed, and democratic" which corresponds to the IRB's rules that include public scrutiny and competence in research areas being reviewed. Likewise, the PP's expectation that decision making "must include affected parties" relates to the Common Rule's mandate to include nonaffiliated and nonscientist members, particularly "members who represent disadvantaged or vulnerable populations from which research subjects may be selected." What remains controversial, however, is the type and extent of remedy that is available to a worried IRB. Basically, the IRB's options are to deny a protocol or request a change in research design or the informed consent statement.

The PP has elicited a large and growing literature and has become influential among policy makers, especially in Europe, and most clearly regarding the environment. Studies of decision making have found that "scientists across the policy spectrum accept and apply some variation of the principle of precaution" in interpreting risk to humans (Silva, Jenkins-Smith, and Barke, 2007). On the other hand, Pielke and Sarewitz (2002) have argued that

environmental researchers, particularly on global climate change, have exploited scientific uncertainty to advance their personal and disciplinary research agendas but to the detriment of informed policy making.

The Precautionary Principle is in some ways in conflict with cost-benefit or risk-benefit comparison or analysis, which is required by the Common Rule. For example, Sunstein (2005) argued that regulatory steps derived from the PP are likely to introduce unforeseen new hazards, rendering the principle self-defeating. At the same time, the specifications required by formal cost-benefit analysis are usually unavailable. Gardiner (2006) offered a variant of the PP building on the Rawlsian maximin principle where uncertainty is large, when some outcomes are simply unacceptable, and when the losses from following the maximin or PP are not important to the decision maker. This situation describes the IRB's task: facing large uncertainties and the possibility of a catastrophic failure such as a lawsuit, negative publicity, suspension of federal funding, or even a subject's death, IRBs are likely to choose a course which will have the least worst of the worst possible outcomes. But to evaluate Gardiner's third criterion -- are IRB decision makers indifferent to the potential losses from being too aggressive in preventing speculated harm? -- requires a consideration of what those losses might be, and how the IRB assesses the benefits from research that might be forestalled by its decisions.

ASSESSING SCIENTIFIC MERIT

In human subjects research, there are four possible components to scientific merit. First, it must comply with standards of research practice within the field. Second, the research might propose a direct benefit to the research subjects, which is usually considered only in clinical research; this topic has been examined in detail by medical ethicists (Freedman, 1987; Weijer and Miller, 2004). Third, the research can promise to advance scientific knowledge, or fourth, to improve society. The fourth aspect is supposedly off limits to the IRB because of the Common Rule prohibition cited above, but we will see that this limitation may be difficult to observe in practice. I will focus primarily on controversies about the IRB's review of scientific methodology and benefits to scientific knowledge.

Methodology

The government guidebook for IRBs offers a convoluted prescription for considering the scientific quality of a proposal:

> "... if a research study is so methodologically flawed that little or no reliable information will result, it is unethical to put subjects at risk or even to inconvenience them through participation in such a study. One question that every IRB member asks is 'To what degree is it our responsibility to review the underlying science of the proposed research?' *Clearly, if it is not good science, it is not ethical*. The federal regulations under which IRBs operate, *however*, do not clearly call for IRB review of the scientific validity of the research design. *Nonetheless*, they do require that IRBs determine whether "[r]isks to subjects are reasonable in relation to...the importance of the knowledge that may reasonably be expected to result" (OHRP, 1993, ch. 4, emphasis added)

International accords on human subjects research state that "scientifically unsound research on human subjects is ipso facto unethical in that it may expose subjects to risks or inconvenience to no purpose" (CIOMS, 1993).

Among the chief concerns about judging research methodology is the disparate expertise of IRB members (whether disciplinary experts or laypersons) across many scientific fields and the role of prior peer review. As discussed below, IRBs would prefer to delegate their review of scientific methodology to preceding peer review processes (Levine, 1986) or departmental review committees, although there is no regular mechanism by which IRBs learn the review history of a proposal. But in July 2003, OHRP stated:

> "In order for the IRB to make the determinations required under HHS regulations . . . the *IRB must receive and thoroughly evaluate sufficient information describing the research design* . . . Furthermore, making the determinations required under HHS regulations at 45 CFR 46.111 *cannot be deferred or*

delegated by the responsible IRB . . . to any other committee or body" (emphasis added; http://www.hhs.gov/ohrp/compliance/letters/2003.html).

The IRB guidebook states that IRB members should "understand the basic features of experimental design," but "not hesitate to consult experts when aspects of research design seem to pose a significant problem" (OHRP, 1993, ch. 4). About half of IRBs report using one or more consultants (NIH, 1998, 38). IRBs occasionally raise questions about research design issues such as criteria for inclusion of participants, threats to validity, and sampling and questionnaire design (Weinberg and Kleinman, 2003). Many IRB members would prefer not to focus on methodological issues, and many in fact do not,[ix] but how much of this they do, one IRB member reported, depends on the "constraints of the local political architecture."

If an IRB finds a significant methodological threat to scientific merit, it might convey these concerns to the researcher, it could approve the proposal conditionally pending a revision of the research protocol return a protocol to the investigator and ask for clarifications, or it might invite the researcher to meet with IRB members to discuss the research plan. A study of the approvals of protocols "as submitted" found that 34 percent of IRBs approved no protocols unchanged, while 10 percent approved only between one-quarter and one-half, and only 6 percent approved more than one-half of protocols as submitted. The problems mostly related to overly technical language in informed consent statements (60 percent), while questions about research design were noted "often" (3 percent) or "sometimes" (58 percent) (NIH, 1998, pp. 61-62). In another study fifty-five percent of IRB members felt that "the scientific quality of research done on human subjects is improved by IRB review" -- but only thirty-seven percent of investigators agreed (NIH, 1998, p. 59).

Of course, a fundamental cross-disciplinary standard by which methodology can be assessed would help IRBs immensely. Empirical studies of how scientists make judgments about science present a complex picture (e.g., Klahr and Simon, 2001; Dunbar, 1999). Nersessian (1999) has employed "model-based reasoning" to suggest that scientists use a variety of types of representations, including thought experiments, visual representations, and analogies, particularly in situations rife with complexity and uncertainty. These result in tacit perceptions and criteria that raise questions about perfect objectivity in the assessment of scientific methods (cf. Shapin and Schaffer, 1985; Kitcher, 2001). Uncertainty in scientific judgments with which IRBs must grapple remains a matter of controversy both to philosophers of science and to policy makers.

"Scientific Benefit"

The benefits of scientific research are complex and difficult to predict. Nevertheless, under IRB guidelines scientific benefit must be addressed. The Common Rule refers to both "anticipated benefits, if any, to subjects," and to "the importance of the knowledge that may reasonably be expected to result" (45 CFR 46.111a(2)). One survey found that although nearly all IRBs take both types of benefit into account, "benefit assessment is not well conceptualized and is often ad hoc, rather than standardized and systematic" (Churchill 2003). Even if the risks are minimal, "examples of research that would not be socially or scientifically valuable include clinical research with nongeneralizable results, a trifling hypothesis, or . . . with results unlikely to be disseminated or in which the intervention could never be practically implemented even if effective."(Emanuel et al., 2000).

Issues of scientific benefit usually are of more direct concern in clinical practice research than in social and behavioral research because the risks to human subjects are likely to be more immediate (but not necessarily more severe: significant potential harm may accompany social and behavioral studies that involve legal risk, psychological impacts, or social effects). For example, in clinical research that compares therapies, "clinical equipoise" requires that "there must be controversy within the scientific community about whether the new intervention is better than standard therapy, including placebo" (Freedman, 1987). Weijer and Miller (2004) have argued that IRBs should judge research concerning therapeutic and nontherapeutic procedures differently because the latter requires consideration of scientific value, not just benefit to the human subject, raising questions of social priorities and community values.

IRBs also must consider the risks to science and society inherent in foregone benefits: an overly cautious treatment of uncertainty about benefits may eliminate the possibility of significant scientific advances (a counterfactual

assessment that would difficult to perform). The Belmont Report began with the assumption that human experimentation is a moral good, perhaps even a moral imperative, given its potential for advancing the general welfare. On the other hand, one of its authors argued that "a slower progress in the conquest of disease would not threaten society, [but] society would indeed be threatened by the erosion of those moral values whose loss, possibly caused by too ruthless a pursuit of scientific progress, would make its most dazzling triumphs not worth having" (Jonas, 1969). Without a benefit-side analogue to the Belmont Report's guidelines on assessing risk, the IRB process is unbalanced.

The Common Rule instructs IRBs to consider "the importance of the knowledge that may reasonably be expected to result," but "the IRB should not consider possible long-range effects of applying knowledge gained in the research." This can be a difficult distinction because although the research process itself might not be harmful to human subjects, the validity of a research conclusion, and therefore the quality of a research design, can be related to how it will be generalized. For example, it might be reasonable to restrict a study of physiological responses to a sports drink to male subjects because of issues related to water retention during women's monthly cycles, but it is unlikely that the beverage would be likely to be marketed only to males. Should an IRB consider this threat to the external validity of the study to be a flaw in research design or a violation of the "justice" provision of the Belmont Report?

Especially when science becomes applied, or when it is used to explain complex systems, the carefully controlled parameters of the laboratory disappear and an undetermined number of possible variables can apply. In the end, research proceeds by a decision to take a chance on scientific benefits – as bioethicist Nancy King has said, "to put hope in its proper context." How that chance is addressed depends largely on who is doing the assessment.

The established practice for evaluating science is peer review. The US Supreme Court has endorsed peer review as an important consideration for judging scientific expertise,[x] and the Department of Health and Human Services has stated that "If data and analytic results have been subjected to formal, independent, external peer review, the information may generally be presumed to be of acceptable objectivity" (US Department of HHS, 2002). However, the peer review process has come under steady criticism. Empirical studies have raised doubts about the effect of peer review on research quality (e.g., Callahan et al., 2002; Jefferson et al., 2002; Goldman and Katz, 1982). Yet many IRB members indicate that they are willing, even eager (given their workload and specialized expertise) to defer on methodological issues to a funding agency's peer review panel (US Department of HHS, 2002, chap. 4).

Peer review does not occur only in granting agencies or by scientific experts. Many universities, hospitals, and other research institutions have separate research committees, especially in areas such as cancer or pediatric research, that review proposals prior to their submission to IRBs or sponsors, assessing their scientific merit and quality, safety, and efficacy. For some areas of research, especially in biomedical research, internal research committees are mandated by funding agencies. When these exist, IRBs are less likely to delve into specific research design questions. Many institutions require department chairs to sign off on proposals before submission to the IRB, but the degree of their scrutiny is uncertain. Several US funding agencies have stipulated that "consumers" be placed on peer review panels (for example, requiring that breast cancer survivors be included on such panels for funding breast cancer research).[xi] In some cases they are credited with recognizing that some proposals may be scientifically intriguing but of little real benefit because "no one in their right mind would undergo the treatment being tested" (Agnew, 1999). But industry-funded drug and device trials do not receive independent peer review, at least not until an FDA application is submitted, and a large proportion of social and behavioral research projects are never submitted for funding, so the IRB may be the only review. In short, the peer review process that might precede IRB review of scientific benefit is far from consistent and uniform.

Efforts to expand peer review in granting agencies to include extra-scientific considerations have not been encouraging. In 1998 the National Science Foundation changed its list of review criteria from four (research merit, relevance, investigators' ability, and impact on science) to two (scientific quality and social impact), with the aim of increasing the role of social impact (primarily education and training, diversity, and relevance to national priorities) in peer reviews. An early study of this change found that "asking scientists to speculate on the possible future broader or societal impacts of a proposal raises a distinct level of discomfort for many reviewers" (NAPA, 2001, 8). Reviewers were likely to use scientific quality as a filter, and only if a proposal passed that test would they consider societal impact. It is far from clear that an assessment of scientific methodology or benefit that is performed separately from an assessment of risk would address all of the methodological issues that might relate to the

risk/benefit comparison required of IRBs; a study of animal research committees observed "it is not clear to us how one does a partial scientific merit review" (Mann and Prentice, 2004).

Furthermore, there are differences in the level of methodological detail and attention to scientific benefit in the requirements for IRB review and a proposal submitted to a granting agency such as NIH or NSF. A research plan submitted to a funding agency must provide sufficient detail to allow judgment on the overall soundness of the research design and the competencies of the investigators, but it often does not describe the specific types of interaction with research subjects (Barke, 2002; but see Kelly and Johnson, 2005).[xii] Therefore, the IRB is likely to have a different standard for judging the quality of scientific methodology than a conventional peer review panel. As one IRB administrator put it, "I don't trust peer review: they look at only the parts of the proposals that they know best, and they miss the big picture." The task of most peer review panels is only to decide whether the research is worthy of funding, not whether it will be done. Under the Common Rule, the IRB cannot escape that responsibility.

BALANCING UNCERTAINTIES

The IRB is predisposed to focus on risk; that is its reason for existence, and in this task it is guided by both broad principles (Belmont) and some specific indicators (e.g., informed consent forms). It has no similar predisposition to focus on scientific benefit, about which its mandate is much less clear and somewhat contradictory, and for which there is no framework for predicting or assessing the gains. Risk is related to magnitude of harm and probability, often with at least a rough estimate of the latter. In contrast, benefit is conceived as a possibility of gain, but often lacks even an implicit estimate of probability. Without some consistent treatment of probability or uncertainty, a comparison of the risks and benefits is itself risky (Churchill et al., 2003).

The idea of a risk-benefit balance in research is persistent but it masks the important role that uncertainty or chance plays in scientific research (Van Ness, 2001). There are many forms of scientific ambiguity, and its treatment depends somewhat on scientists' individual preferences (Anand, 2002). Daniel Sarewitz has written that "scientific uncertainty. . . can be understood not as a lack of scientific understanding but as the lack of coherence among competing scientific understandings, amplified by the various political, cultural, and institutional contexts within which science is carried out" (2004, 385). Certainly in the case of IRB reviews of speculative research proposals some "lack of scientific understanding" comprises at least some of the scientific uncertainty, but the other factors mentioned by Sarewitz also certainly play a complicating role in the anticipation of both risk and benefit.

The IRB has both a scientific task and a regulatory task as it polices human subjects research. Scientific communities and legal systems consider risk differently, with the burdens of proof used in science to establish a causal relationship tending to be much more rigorous than those used in legal reasoning (Cranor, 1993). Similarly, "the considerations that scientists bring to bear when evaluating scientific evidence about risk seldom appear in coverage of health issues by the lay media; it is as if scientists and the lay public inhabit different worlds entirely" (Foster et al., 1993). It is unclear whether lay members of the IRB interpret uncertainty differently from scientists, in part because they may be socialized into the norms of scientific deference after a few IRB meetings (Lane, 2005), but even the scientists on an IRB may be pulled in several directions by the IRB's mandate. This diversity was designed into the IRB by the Common Rule's mandates on membership.

Balancing and finding ways past these uncertainties requires balancing the IRB's responsibilities: externally, to the community and its citizens, and internally, to the advancement of scientific research, but without bias toward the interests of the IRB's research institution. The different roles and missions of the various players in the research enterprise are a reminder that science is a system of individuals and institutions (not all of them scientist or scientific) that interact in ways that constrain each other. Government agencies have missions to protect the public health and advance knowledge. Research sponsors want to get new products to publication or market. Research institutions promote science, technology, and medical care while promoting their financial gain and avoiding costly errors. Researchers want to explore, educate, commercialize, publish, take risks, minimize errors, and avoid hurting anyone. And human research subjects are very much a part of the enterprise, with responsibility to hold researchers, and themselves, to a high standard when they consider an informed consent statement.

A consistent solution to the incommensurability of uncertainty in assessing risks and benefits in IRBs is not only unlikely, but also possibly undesirable. "When facts and values overlap, and are deeply controversial, the only opportunity for mutual understanding may be to look for practical, 'local' answers, based on different positional insights" (Pellizoni, 2003).[xiii] Similarly, arguments about solutions to incommensurability and noncomparability have noted that balancing is shaped by factors such as "zones of indeterminancy" and the "absence of covering values" (Chang, 1997), which relate well to the idea that uncertainty can facilitate local compromises based on diverse individual and institutional preferences.

The practice of risk assessment blends scientific risk analysis and ethics. Writing of the shared values of scientists, Kuhn wrote that "Judgments of accuracy are relatively, though not entirely, stable from one time to another and from one member to another . . . But judgments of simplicity, consistency, plausibility, and so on often vary greatly from individual to individual" (1972, 185). At national workshops IRB members report that in the face of uncertainty their risk judgments ("should it be done?") often become narratives about "would I let my sister, or father, or child, participate in this project?" The analogous filter or question for judging scientific merit is for scientist members to ask themselves "would I do it this way?" rather than "would I do this at all?"

A strict comparison of risks and benefits also would require that the assessment of risks and benefits be independent of each other. But researchers make many research design decisions based on the tradeoffs between scientific benefit (especially statistical significance and generalizability) and risks to human subjects (how many, and from which subpopulations) (Bacchetti et al., 2005). High methodological quality and concern for research ethics are not antithetical. A study of published clinical trial research found that "trials with higher standards of quality met the requirement of reporting ethical issues more frequently" than those studies assessed as lower in methodological quality (Ruiz-Canela et al., 2001, 174). Cognitively, Slovic (2002) has found evidence of an "affect heuristic," in which people have an initial emotional reaction to a situation which then "directs" their judgments of both risk and benefit.

The particular expertise of IRB members plays an important role in IRB decision making. There are at least three types of expertise on the IRB. First, there is the epistemic expert: the experienced researcher who is familiar with the methodology and context of a particular research protocol -- for example, a sociologist reviewing a protocol concerning a survey of high school students' attitudes on alcohol abuse. Second, there is the researcher who is knowledgeable about science and its processes but expert in a field other than that of the protocol being reviewed, such as a chemist who is reviewing the alcohol use study. Third, there is the non-expert, presumably a nonscientist who often also serves on the IRB to represent the community or particular vulnerable populations. These often are "experience based experts": "members of the public who have special technical expertise in virtue of experience that is not recognized by degrees or other certificates" (Collins and Evans, 2002, 238). However, the confidence of scientists in commenting on risk and ethics is not balanced by the confidence of non-scientists on the IRB to assess science.

One way to consider the challenge faced by IRBs in balancing uncertainties is to divide research protocols into four groups based on the degree of risk to human subjects and on the methodological quality of the research design. First, most protocols are well designed and impose very little risk; for these, comparisons of risks and benefits are unnecessary. Other protocols are well designed but include minor risks that can be addressed by research investigators, sometimes in response to questions by the IRB. Another category includes clearly poorly designed projects posing significant risks to human subjects; this is almost a null set for IRBs. A fourth category is more challenging: research protocols that impose minimal risk but also include methodological flaws. Experts can find design problems in almost any research proposal (after all, peer reviewers seldom give proposals a perfect pass), so for these protocols IRBs must decide whether to advise researchers of possible flaws in the research design, including those that have no effect on risks to human subjects. Even if an IRB adheres to the "bad research cannot produce good results" perspective, its obligations in such cases are not clear, and many IRBs, researchers, and experts on human subjects research disagree strongly about how far to intrude on researchers' autonomy.[xiv]

A review of research design would be seen by many expert IRB members as largely objective, based on disciplinary norms -- is the sample size large enough, are the questions worded clearly, etc. -- while the assessment of risk is inherently messy, wrestling with speculated harms and human subjects' own perceptions of risk. IRBs' tendencies to focus on the structure and wording of the informed consent statement may be an attempt to pull ethical questions

into the realm of formulaic prose and structured risk. After all, with a well-constructed informed consent statement, much of the burden of assessing the acceptability of risk is transferred to the IRB research subjects who presumably should judge the risks for themselves. Part of the myth of science is that it allegedly embodies a closed and knowable set of processes for advancing knowledge, while ethics has no right parenthesis. Without improved and clearer guidelines, the IRB system will continue to be inconsistent in the consideration of risks and benefits.

The Sanguinity Principle

The precautionary principle focuses on the possibility of adverse events. There is no corresponding articulated principle for assessing the possibility of scientific benefits. "Science" for an IRB encompasses both basic research undertaken with no expectation of immediate social or economic benefit, as well as applied research undertaken with specific benefits in mind, ranging from market opportunities to policy inputs. Scientific benefits also range from the proximate (such as contributions to scholarly literature, the training of students, patentable innovations) to the long-term (e.g., the education of future generations, contributions to a pool of theoretical knowledge, paradigm shifts) The Common Rule specifies (§46.111) that the IRB consider not only risks to human subjects but also "the importance of the knowledge that may reasonably be expected to result," but although the law also states that IRBs "should not consider possible long-range effects of applying knowledge gained in the research (for example, the possible effects of the research on public policy) as among those research *risks* that fall within the purview of its responsibility" (emphasis added), it does not similarly limit the IRB's assessment of scientific benefits to short-term effects, nor does it define "long-range" (Casarett, Karlawish, and Moreno, 2002). Therefore an IRB has considerable discretion in defining scientific benefits, and a guiding principle for assessing scientific benefits, so broadly definable, must itself be broad.

IRB members who are also researchers are familiar with the problem of anticipating results: is the study well-formulated with a coherent hypothesis or research question, will useful data be available, are there adequate controls for exogenous factors, is there enough time and money – and will the results be publishable, published, and perhaps patented (before someone else addresses the same topic sooner and better)? With this cascade of contingencies, an expectation of benefits is more likely to be based on the general desirability of conducting good research than on a specific outcome from a specific research project. The scale is tilted toward a tacit faith that scientific research is worthwhile, with the benefits mostly unknown, unknowable, and therefore usually not thoroughly scrutinized in IRB reviews. As discussed, many proposals face no more than departmental review (often in the form of a signature from a department chair); in any case, peer review processes at NSF and NIH apply different standards, including the level of methodological detail and whether risks and benefits are assessed. From her interviews with IRB members Lane (2005) found that some IRBs report that they do not assess scientific merit (conflating methodology and outcomes) at all, while in those IRBs that do judge scientific benefit IRB scientist-members reviewing proposals from a particular discipline tend to defer to other members with corresponding disciplinary expertise. Such variability leads to the question of what, if anything, holds IRBs together in performing their task of considering the benefits of research.

Just as it would be inaccurate to synonymize precaution with pessimism – the concern is not that the worst *will* happen, but that it *might* -- the approach to scientific benefit corresponds not to blind optimism but to sanguinity: a "disposition toward hopefulness or confidence of success." The risk-side precautionary principle has a broad benefit-side counterpart: a *sanguinity principle* that endorses research despite uncertainty about the benefits of a project. Such a principle promotes the cause of science as a natural good, it denies the necessity (or possibility) of a rigorous risk-benefit analysis to evaluate an individual research project, it puts the burden of proof on those who would reject the possibility that a project will advance knowledge, it embraces a willingness to entertain untested but reasonably plausible research questions, and it tolerates research that proposes only incremental benefits. One need not subscribe fully to the notion of paradigms and their shifts to recognize that Kuhn's interpretation of normal science, within which much of the work consists of necessary but incremental "mopping-up," relies heavily on the inclination of scientists to take a generally positive view of the current state of research in their field. As Kuhn wrote, scientists can "agree in their *identification* of a paradigm without agreeing on, or even attempting to produce a full *interpretation* or *rationalization* of it" (Kuhn, 1970, 44; emphasis in original).

Both scientists and large majorities in the lay public are inclined to believe that doing science is itself beneficial, and that science generally brings benefits to society. Most scientists would identify with the 81 percent of public survey respondents in the US who agreed that: "Even if it brings no immediate benefits, scientific research that advances the frontiers of knowledge is necessary and should be supported by the Federal Government," and with the 72 percent who said that the benefits of science strongly or slightly outweighed its harms[xv] (NSF 2001, ch. 7). This disposition toward optimism about inquiry permeates research institutions and the research enterprise IRBs regulate. Even the nonaffiliated non-expert members of the IRB are unlikely to be cynical about the overall utility of science, or they would probably not be volunteer members.[xvi]

What issues about a research protocol could cause IRB members to discard their sanguinity? It could be doubts about the qualifications of the researchers, the structure of the research design, the project's relevance to existing knowledge and practice, or other matters. However, in some cases what prompts a question about scientific benefit is the nature of the *risk*: if the risk is unmistakably unreasonable (unnecessarily non-minimal, avoidable, or far out of proportion to purported benefits), the IRB members are likely to express doubts about the quality of thinking that went into the science. If the assessment of risk cannot be neatly isolated from the assessment of scientific benefit, then a strict comparison or a meaningful risk-benefit ratio is much more difficult.

CONCLUSION

As a guide for IRB decision making, the Belmont Report and Common Rule provide more of an exhortation than a template: do no harm, at least not unless the harm is outweighed by the benefits. But how is this calculus to be implemented? Given their mandate and institutional environments, IRBs are expected to approach their task in three ways simultaneously: as primarily ethical (agents of equity), scientific (quasi-dispassionate assessors of risks and benefits), or bureaucratic (move the protocols through the system). How they handle the various uncertainties that are endemic to these tasks is crucial to understanding how they balance these responsibilities.

In the case of the IRB uncertainty may serve as a mediating device that allows bridging of experiences and perspectives. In considering uncertainty, natural scientists tend to focus more on ontological limitations such as problems in measurement and replicability (e.g., Hund et al., 2001), while social scientists are more concerned with stochastic uncertainties (i.e., the inherent unpredictability of human and social systems) or epistemic uncertainties deriving from subjectivities in knowledge and perceptions (Marris et al., 1998; Pate-Cornell, 2002). The epistemologies and the discourses used to interpret uncertainty and apply it to science-based decision making have been shown to differ across disciplines (Romanello, 2003; Barke and Jenkins-Smith, 1993). Yet uncertainty applies to all aspects of IRB consideration, and because the type of science considered varies widely on many IRBs (particularly those that consider both biomedical and behavioral/social science protocols), different uncertainties about risk, benefits, and other aspects of research protocols are likely to be voiced and shared by each IRB member. Precaution and sanguinity are useful principles for negotiating uncertain risks and benefits in such situations. Scholars have argued that uncertainty, whether about risk or about science, is often used as a boundary-ordering device that "allows scientists (i) to translate uncertainty for policymakers so as to make its reduction appear more tractable and (ii) to maintain a richer, or more heterogeneous, version of uncertainty for scientific communities than for policymakers" (Shackley and Wynne, 1996, 293; Zehr, 2000). In the case of the IRB, precautionary and sanguinary approaches help IRB members to find tacit commensurability between different types of uncertainty, creating a consensual negotiating space that allows decisions.

This intertwining of judgments of risks and benefits, and the possibilities for compromise – among members, and between ethics and science -- that is created by uncertainty about both, has implications for recommendations that scientific assessments and ethical review, now conjoined in IRBs and research ethics committees, should be separated. A study by the UK Department of Health in 2005 (the "Lord Warner" report) concluded that "RECs should deal with ethical rather than scientific review" (UK Department of Health, 2005, 13). Similarly, a 2002 report by the US Institute of Medicine recommended a distinct, three-way allocation of responsibilities for protecting human research participants: scientific and financial conflict of interest reviews each would be conducted prior to and separately from ethics review, which would be conducted by a "research ethics review board" (or "Research ERB," no longer an "IRB"). The report found that "the scientific and ethical review of protocols should be equally rigorous. Therefore, each review requires distinct, although overlapping, expertise"

because "it is unrealistic to expect a single group of individuals to possess the requisite skills to competently carry out" these tasks (IOM, 2002, 10, 72). Permitting an IRB to conduct a scientific review may distract it from a thorough review of ethical issues, or it "may lack the scientific expertise necessary to adequately assess the technical merit of a proposal" (76). However, "in rare and controversial cases" such as stem cell or xenotransplantation research "it may be appropriate to pursue the ethical consideration of a protocol before, or in conjunction with, the evaluation of its scientific merit" (77).

But the ease of separating science review from ethics review may be undesirable (Epstein, 1996) and dangerous to assume. The intertwining of ethics and research design is illustrated by the concept of external validity. The generalizability of research findings to larger populations is a fundamental issue in research design: it depends on the size and selection of the study sample, but also vice versa. A research design embodies decisions about the intended impact of the study, and the inclusion or exclusion of particular subpopulations is inherently an ethical issue. For example, in a regression discontinuity research design, test subjects are placed into a group based on a particular criterion often related to vulnerability or need. This design is used when the researcher "intend[s] to balance ethical and scientific concerns when it is deemed unethical or infeasible to randomize all patients into study treatments" (Trochim and Cappelleri, 1992, p. 387). Such an approach carries ethical implications related to the number of patients that would be denied a beneficial drug or exposed to a harmful drug. Thus, it is not clear that the scientific and ethical issues are as separable as the Lord Warner and IOM reports assert. The Environmental Protection Agency's Science Advisory Board found that "The separation of science and ethics -- as occurs when scientific peer review precedes the evaluation of a study by an Institutional Review Board (IRB) -- may be procedurally necessary, but it is a separation that is arbitrary and difficult to defend. . ." (EPA, 2000, 24).

IRB behavior must be understood in the context of a larger system of science institutions and actors. With its role in this system, the IRB's use of a precautionary principle in assessing risks to human subjects is appropriate; it matches the mandates of the Belmont Report and the Common Rule, and it allows action in the face of inadequate information about the interests of research subjects. To balance, the sanguinity principle is consistent with the aims of researchers and their institutions given the uncertainties in science and the absence of a framework for prediction of scientific benefit.

In science, where empirical knowledge is inherently uncertain and all theories are subject to refinement or refutation, uncertainty plays a different role than in clinical research. In the latter, uncertainty can directly affect the lives of patients or human subjects. Medical ethicists have many approaches to coping with such uncertainty, but they generally break down into questions about how much uncertainty about benefits and risks can be tolerated by patients, physicians, and the medical community. Similar questions are not applied to judging the risks and benefits of most non-clinical scientific research. Further examination is needed of differences in the assessment of benefit and risk between scientific and clinical research: for example, what are appropriate IRB standards for evaluating clinical research protocols that do not aspire to scientific generalizability? We also need to understand the applicability of medical ethics principles such as "clinical/community equipoise" and "the uncertainty principle" to scientific research outside the biomedical realm.[xvii]

Because IRBs are composed of an uncommon combination of insider and outsider experts and non-experts as members, and serve a mission of balancing imprecise objectives intended to protect both scientific advancement and the participants in scientific studies, the IRB is a peculiar institution within the science policy process where the normal forces and authorities of science might be expected to be less likely to apply. Yet they are laboratories in which we can examine the operation of what has been proposed for decades: direct involvement in determining the priorities and techniques of science by non-scientists and scientists outside a normal narrow disciplinary community. Questions about how risks and benefits, ethics and science, and knowledge and uncertainty are reconciled by these diverse committees can shed light on many important questions in science policy.

AUTHOR'S NOTE: Several colleagues provided helpful comments on this paper: Jilda Garton, Roberta Berry, Jason Borenstein, and Eliesh Lane, as well as several anonymous reviewers. An earlier version of this paper was presented at the 2005 meeting of the American Association for the Advancement of Science.

REFERENCES

Agnew, Bruce. 1999. "NIH Invites Activists Into the Inner Sanctum," *Science* 283, March 26, 1999-2001.

Anand, P. 2002. "Decision-Making When Science is Ambiguous," *Science*, 295 (8 March), 1839.

Ashcroft, Richard, et al. 2005. "Reforming Research Ethics Committees," *British Medical Journal*, 331 (Sept. 17), 587-588.

Bacchetti. P., L.E. Wolf, M.R. Segal, and C.E. McCulloch. 2005. "Ethics and Sample Size," *American Journal of Epidemiology* 161, 105-110.

Barke, Richard P. 2002. "Authorities of Knowledge in Research Centers: The Role of Research Ethics Committees in Peer Review," paper presented at World Congress of Sociology, Brisbane, Australia, 7-13 July.

Barke, Richard P., and Hank C. Jenkins-Smith. 1993. "Politics and Scientific Expertise: Scientists, Risk Perception, and Nuclear Waste Policy," *Risk Analysis*, 13, 425-439.

Callahan, M. L., R. K. Knopp, and E. J. Gallagher. 2002. "Effect of Written Feedback by Editors on Quality of Reviews: Two Randomized Trials," *Journal of the American Medical Association* 287, June 5, 2781-2784.

Casarett, David J., J. Karlawish, and J. Moreno. 2002. "A Taxonomy of Value in Clinical Research," *IRB: Ethics and Human Research* 24 (Nov.-Dec.), 1-6.

Chang, Ruth, ed. 1997. *Incommensurability, Incomparability, and Practical Reason.* Harvard Univ. Press.

Churchill, Larry R. et al. 2003. "Assessing Benefits in Clinical Research: Why Diversity in Benefit Assessment Can Be Risky," *IRB: Ethics & Human Research*.

Collins, Harry M. and Evans, R. 2002. "The Third Wave of Science Studies: Studies of Expertise and Experience," *Social Studies of Science*, 32(2): 235-296, April.

Collins, Harry and Trevor Pinch. 2005. *Dr. Golem: How to Think about Medicine*. Univ of Chicago Press.

Council for International Organizations of Medical Sciences. 2003. *International Ethical Guidelines for Biomedical Research Involving Human Subjects*. Geneva, Switzerland: CIOMS.

Covello, V. 1992. "Risk communication: An emerging area of health communication research," in *Communication Yearbook* 15 (S. Deetz, ed.) Newbury Park and London: Sage, 359-373.

Cranor, Carl F. 1993. *Regulating Toxic Substances: A Philosophy of Science and the Law*. New York: Oxford University Press.

Douglas, M. and Wildavsky, A. 1982. *Risk and Culture*. Berkeley: University of California Press.

Dunbar, K. 1999. "How scientists build models invivo science as a window on the scientific mind," In L. Magnani, N. J. Nersessian, & P. Thagard (eds.), *Model-based reasoning in scientific discovery*. New York: Kluwer, 85-99.

Emanuel, Ezekiel J., David Wendler, and Christine Grady. 2000. "What Makes Clinical Research Ethical?" *JAMA* 283 (20), May 24/31, 2701-11.

Epstein, Steve. 1996. *Impure Science: AIDS, Activism, and the Politics of Knowledge.* Berkeley, CA: University of California Press.

Foster, Kenneth R., David E. Bernstein, and Peter W. Huber. 1993. *Phantom Risk: Scientific Inference and the Law*. Cambridge: MIT Press.

Freedman, B. 1987. "Equipoise and the ethics of clinical research," *New England Journal of Medicine*. 317:141-145.

Gardiner, Stephen. 2006. "A Core Precautionary Principle." *Journal of Political Philosophy* 14 (March) 33-60..

Goldman, J. and M. D. Katz. 1982. "Inconsistency and Institutional Review Boards," Vol. 248 No. 2, July 9.

Goldstein, Bernard D. 1999. "The Precautionary Principle and Scientific Research Are Not Antithetical," *Environmental Health Perspective*s, 107 (12), December.

Groth, E. 1991. "Communicating with consumers about food safety and risk issues." *Food Technology* 45(5): 248-253.

Gunsalus, C. Kristina et al. 2006. "The Illinois White Paper - Improving the System for Protecting Human Subjects: Counteracting IRB Mission Creep," U Illinois Law & Economics Research Paper No. LE06-016 Available at SSRN: http://ssrn.com/abstract=902995

Hamburger, Philip. 2005. "The New Censorship: Institutional Review Boards," *Supreme Court Review*, 271-354. Available at SSRN: http://ssrn.com/abstract=721363.

Hund, E., D. Massart, and J. Smeyers-Verbecke. 2001. "Operational Definitions of Uncertainty," *Trends in Analytical Chemistry* 20: 394-406.

Institute of Medicine (2002) *Responsible Research: A Systems Approach to Protecting Research Participants*, Federman, D., Hanna, K., Rodriguez, L. (Eds), Washington, DC: The National Academies Press.

Jefferson, T., P. Alderson, E. Wager, and F. Davidoff. 2002. "Effects of Editorial Peer Review: A Systematic Review," *Journal of the American Medical Association* 287, June 5: 2784-2786.

Jonas, H. 1969. "Philosophical Reflections on Experimenting with Human Subjects." *Daedalus* 98: 219–247.

Jonsen, Albert R., and Stephen Toulmin. 1988. *The Abuse of Casuistry: A History of Moral Reasoning*. Berkeley: University of California Press.

Keiger, Dale and Sue De Pasquale, S. 2002. "Trials & Tribulations," *John Hopkins Magazine*, 54 (1), February, at http://www.jhu.edu/~jhumag/0202web/

Kitcher, Philip. 2001. *Science, Truth, and Democracy*. Oxford Univ. Press.

Klahr, D., & Simon, H. A. 2001. "What have psychologists (and others) discovered about the process of scientific discovery?" *Current Directions in Psychological Science*, 10: 75-79.

Kuhn, Thomas. 1972. *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Lane, Eliesh O'Neil. 2005. *Decision-Making in the Human Subjects Review System*, Atlanta: Georgia Institute of Technology.

Levine, Robert J. 1986. *Ethics and Regulation of Clinical Research*. 2d ed. Baltimore: Urban and Schwarzenberg.

Mann, Michael D., and Ernest D. Prentice. 2004. "Should IACUCs Review Scientific Merit of Animal Research Projects?", *Lab Animal* 33 (January): 26-31.

Marris, C., I.H. Langford, and T. O'Riordan. 1998. "A Quantitative Test of the Cultural Theory of Risk Perceptions: Comparison with the Psychometric Paradigm," *Risk Analysis* 18: 635-647.

Martinson, Brian C., Melissa S. Anderson, and Raymond de Vries. 2005. "Scientists Behaving Badly," *Nature* 435 (June 9): 737-738.

Montague, Peter. 1998. *Rachel's Environment & Health Weekly* #586, 19 February.

National Academy Of Public Administration. 2001. *A Study Of The National Science Foundation's Criteria For Project Selection*, February.

National Institutes of Health. 1998. "Evaluation of NIH Implementation of Section 491 of the Public Health Service Act, Mandating a Program of Protection for Research Subjects," Final Report ("Bell Report"). NIH Office of Extramural Research, June 15.

NRC/NAS Committee on the Institutional Means for Assessment of Risks to Public Health. 1983. *Risk Assessment in the Federal Government*.

National Research Council. 2004. *Intentional Human Dosing Studies for EPA Regulatory Purposes: Scientific and Ethical Issues*, Washington: National Academies Press,

National Science Foundation, 2001. *Science and Engineering Indicators 2000.* Washington: National Science Foundation.

Nersessian, Nancy. 1999. "Model-based Reasoning in Conceptual Change," in Nancy Nersessian, Lorenzo Magnani, and Paul Thagard, eds., *Model-based Reasoning in Scientific Discovery*, New York, Kluwer Academic/Plenum Publishers.

Office of Human Research Protection. 1993. *IRB Guidebook.* Washington: Department of Health and Human Services.

O'Riordan, Tim and James Cameron. 1994. *Interpreting the Precautionary Principle*, Earthscan Publications.

Pate-Cornell, Elisabeth. 2002. "Risk and Uncertainty Analysis in Government Safety Decisions," *Risk Analysis*, 22: 633 -646.

Pellizoni, Luigi. 2003. "Uncertainty and Participatory Democracy" *Environmental Values*, (May) 12: 195-224.

Peto, Richard and Colin Baigent, 1998. "Trials: The Next 50 Years," *British Medical Journal* 317 (Oct. 31), 1170-1171.

Pielke, Roger, Jr. and Daniel Sarewitz. 2002. "Wanted: Scientific Leadership on Climate," *Issues in Science and Technology* (online at http://issues.org/19.2/p_pielke.htm#, last accessed Feb. 2007).

Porter, Joan P. 1986. "What are the Ideal Characteristics of Unaffiliated/Nonscientist IRB Members?" *IRB* 8 (May/June), 1-6.

------. 1987. "How Unaffiliated/Nonscientist Members of Institutional Review Boards See Their Roles," *IRB* 9 (Nov./Dec.), 1-6.

Romanello, S., B. Dervin, and R. Fortner. (2003). "Sense-making as a methodology to inform interdisciplinary communication of scientific uncertainty of global climate change," paper presented at the International Communication Association, San Diego, CA.

Ruiz-Canela, Miguel, et al., 2001. "Methodological Quality and Reporting of Ethical Requirements in Clinical Trials," *Journal of Medical Ethics* 27: 172-176.

Sackett, David L., et al., 2000. *Evidence-Based Medicine*, 2nd ed.. London: Churchill Livingstone.

Sackett, David L. 2001. "Uncertainty about Clinical Equipoise." *Canadian Medical Association Journal*. 164:1831-32

Sandman, P. M. 1987. "Risk Communication: Facing Public Outrage," *EPA Journal* 13: 21.

Sarewitz, Daniel. 2004. "How Science Makes Environmental Controversies Worse," *Environmental Science and Policy* 7 (October), 385-403.

Shackley, Simon, and Brian Wynne. 1996. "Representing uncertainty in global climate change science and policy: boundary-ordering devices and authority." *Science, Technology, and Human Values* 21: 275-302.

Shapin, Steven, and Simon Schaffer. 1985. *Leviathan and the Air-Pump*, Princeton Univ. Press.

Silva, Carol L., Hank C. Jenkins-Smith, and Richard P. Barke. 2007. "The Politics of Caution: The Bases of Scientists' Precautionary Choices," *Risk Analysis*, forthcoming.

Slovic, P. 1987. "Perception of Risk." *Science* 236: 280-285.

Slovic, Paul et al., 2002. "The Affect Heuristic," in Tom Gilovich et al., eds., *Intuitive Judgment: Heuristics and Biases*. Cambridge Univ. Press.

Sunstein, Cass R. 2005. *Laws of Fear: Beyond the Precautionary Principle*. Cambridge Univ. Press.

Trochim, W. and Cappelleri, J.C. 1992. "Cutoff assignment strategies for enhancing randomized clinical trials." *Controlled Clinical Trials*, 13, 190-212.

US Department of Health and Human Services. 2002. *IRB Guidebook*.

Van Ness, Peter H. 2001. "The Concept of Risk in Biomedical Research Involving Human Subjects," *Bioethics*. August; 15: 364-370.

Wald, David S. 2004. "Bureaucracy of Ethics Applications," *British Medical Journal* 329 (July 31), 282-284.

Weick, Karl E. and Kathleen M. Cutcliffe. 2001. *Managing the Unexpected: Assuring High Performance in an Age of Complexity*, Jossey-Bass.

Weijer, Charles and Paul B. Miller. 2004. "When Are Research Risks Reasonable in Relation to Anticipated Benefits?" *Nature Medicine*, June, 10: 570-573.

Weinberg, Janice M. and Ken P. Kleinman. 2003. "Good Study Design and Analysis Plans as Features of Ethical Research with Humans," *IRB: Ethics & Human Research* 25, September-October: 1-14.

Zehr, Stephen C. 2000. "Public Representations of Scientific Uncertainty about Global Climate Change," *Public Understanding of Science* 9: 85-103.

[i] A compilation of human subject research protections in more than 80 nations is available at http://www.hhs.gov/ohrp/international/index.html#NatlPol (last visited Feb. 10, 2007). A list of international institutions that are registered with the US government's Office of Human Research Protections can be found at http://ohrp.cit.nih.gov/search/actrypck.asp (last visited Feb. 10, 2007).

[ii] Hamburger (2005) has argued that IRB regulations comprise an unconstitutional constraint on First Amendment rights by licensing speech and publication.

[iii] I particularly want to avoid the argument extending over many years about whether medicine and clinical research are more science or more art (cf., Sackett, 2000; Collins and Pinch, 2005).

[iv] Research institutions are not neutral actors in the research process. IRBs are required to consider financial conflicts of interest as they might relate to the design of a study or the recruitment of subjects. Institutions are often compensated in industry-funded clinical trials on a per-patient basis, and states have promoted their research environment by advertising the availability of research subjects for pharmaceutical studies.

[v] Many of the observations and conclusions in this article are based on the author's years of experience serving on his university's IRB, as well as discussions and participation at many national IRB conferences and workshops. The conclusions have been presented at national meetings of IRB practitioners and administrators and improved by their comments.

[vi] The World Medical Association's Declaration of Helsinki includes no similar limitation. Principle 16 states "Every medical research project involving human subjects should be preceded by careful assessment of predictable risks and burdens in comparison with foreseeable benefits to the subject or to others."

[vii] "The uncertainties inherent in risk assessment can be grouped in two general categories: missing or ambiguous information on a particular substance and gaps in current scientific theory. When scientific uncertainty is encountered in the risk assessment process, inferential bridges are needed to allow the process to continue. . . . The judgments made by the scientist/risk assessor for each component of risk assessment often entail a choice among several scientifically plausible options; the Committee has designated these inference options" (NRC, 1983: 28).

[viii] Beldsoe et al. (2007) take a slightly different approach, arguing that "the IRB's over-riding goal is clear: to avoid the enormous risk to the institution of being found in noncompliance by OHRP" (p. 14), and they use "the art of pre-emption in high-risk organizations" (Weick and Sutcliffe, 2001) as their lens.

[ix] Lane (2005) found that 71 percent of her scientist-member interviewees favored reviewing research designs, compared with 31 percent of nonscientists.

[x] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[xi] A description of the National Cancer Institute's "Consumer Advocates in Research and Related Activities" (CARRA) program is at http://la.cancer.gov/carra/.

[xii] For most NIH grants (R01), the criteria are (1) significance (does the study address an important problem and will scientific knowledge be advanced?), (2) approach (are the conceptual framework, design, methods, and analyses appropriate to the aims of the project?), (3) innovation (are the aims, concepts, or methods of the project original and innovative?), (4) investigator (is the researcher appropriately trained and experienced?), and (5) environment (does the scientific environment and the institution in which the work will be done contribute to the probability of success?). Peer reviewers applying these criteria are not compelled to ask whether beneficence, equity, and respect for persons (the basic Belmont principles) are observed.

[xiii] An account of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which produced the Belmont Report, says that the commissioners were able to reach agreement by basing their arguments and reasoning on specific cases rather than bioethical theories, and that only when they tried to explain their agreements did they find themselves divided by disciplinary lines (Jonsen and Toulmin, 1988).

[xiv] These issues are occasionally discussed at length in online IRB discussion groups. See, for example, http://www.irbforum.org/forum/read.php?f=3&i=5957&t=5957.

[xv] A 2002 survey of scientists found that 15.5% had changed a research design or results in response to pressure from a funding source, 12.5% had overlooked others' use of flawed data, 7.6% had admitted to "circumventing certain minor aspects of human-subject requirements," and 0.3% had ignored "major aspects" of those requirements (Martinson et al., 2005).

[xvi] A survey of IRB lay members found that 61 percent agreed that their role included "judging the worth of the research," while 34 percent disagreed (Porter, 1987). They also rated "intelligence, judgment, and analytical thinking ability" just behind "self-confidence and courage" as ideal qualities of lay members (Porter, 1986).

[xvii] The uncertainty principle states that patients can be entered into randomized clinical trials only if "the responsible clinician is substantially uncertain which of the trial treatments would be most appropriate for that particular patient. . . in comparison with either no treatment or some other treatment that could be offered to the patient in or outside the trial" (Peto and Baigent, 1998, 1170). The equipoise principle holds that the medical

community must be uncertain about which treatment is better before randomly assigning patients to clinical trials (Freedman, 1987; Weijer and Miller, 2004)