Are IRBs the "least worst" way to promote research ethics?

The Pathologies of Institutional Review Boards

By David A. Hyman
University of Illinois College of Law

Institutional Review Boards (IRBs) are polarizing institutions. IRB supporters view them as the best thing since sliced bread. Detractors believe IRBs impose costs and have no benefits. Supporters point to the good faith and hard work of those who volunteer to serve on an IRB. Detractors suggest that IRBs emphasize bureaucratic busy-work. Supporters ask for more money and more staff so they can do an even more thorough job of reviewing research protocols. Detractors point out that the IRB framework of research oversight would never be approved by an IRB. Supporters counter that notorious examples of abuse (e.g., Tuskegee, Nuremberg) show that IRBs are necessary. Detractors respond with anecdotes of IRB stupidity and incompetence. Supporters argue that conducting research is a privilege, not a right. Detractors complain about censorship, restrictions on academic freedom, and the chilling of constitutionally protected free speech. Both sides then return to their respective camps, secure in the knowledge that they are right and those on the other side are self-righteous zealots.

The controversy over IRBs arises from differing preferences, methodological commitments, and risk tolerances. Both sides believe fundamental principles (academic freedom/censorship vs. the protection of vulnerable human subjects) are at stake, so the dispute is not readily subject to compromise. Even King Solomon would find it difficult to solve the controversy in a way that makes everyone happy — and the original Solomonic strategy (cutting the director of each IRB in half) seems unlikely to improve matters.

David A. Hyman is the Galowich-Huizenga Faculty Scholar and professor of law and medicine at the University of Illinois College of Law. He served on the committee that prepared the AAUP report referenced in this article.

The article is adapted from a forthcoming article that will appear in the *Northwestern University Law Review*.

This article offers some perspective on the dispute, and some modest strategies for improving on the status quo.

WHY DO WE HAVE IRBS?

Federal regulations (the "Common Rule") require all research funded by the federal government and involving human subjects to be overseen by an IRB. Those regulations are rooted in the work of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which was created in response to several well-publicized biomedical research scandals.

"Research" is defined by the regulations as "a systematic investigation... designed to develop or contribute to generalizable knowledge." A human subject is "a living individual about whom an investigator conducting research obtains data through intervention or interaction with the individual, or identifiable private information."

Each IRB must be composed of at least five members, and the members must be from diverse cultural and racial backgrounds. At least one member of the IRB should have scientific expertise, while at least one individual must be a nonscientist. Each IRB must have at least one member who is not otherwise affiliated with the research facility and has no immediate family members who are so affiliated.

An IRB has the authority to approve, require modification of, or disapprove research, both in its initial determination and as part of mandatory continuing (at least yearly) review. In determining whether to approve a study, an IRB is required to evaluate whether the risks to subjects are minimized; whether those risks are reasonable in light of expected benefits; and whether subjects are selected in an equitable manner, with due concern for the particularized risks of conducting research in vulnerable populations. For research that involves multiple

institutions, an institution that obtains the appropriate approvals can "enter into a joint review arrangement, rely upon the review of another qualified IRB, or make similar arrangements for avoiding duplication of effort."

IRBs are also responsible for ensuring that informed consent is obtained from study participants. The regulations specify a wide range of information that must be provided to study participants, including the statement that the study involves research, and a description of the procedures, expected duration, and reasonably foreseeable risks to the subject. An IRB may approve an abbreviated consent procedure or waive consent entirely under limited circumstances.

The IRB can conduct an expedited review in instances in which there is minimal risk from the research. Minimal risk means that "the probability and magnitude of harm or discomfort anticipated in the 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." However, even expedited review can impose a delay of several weeks, and IRBs can always require full-blown approval if they have any concerns about the research.

There are six categories of research exempt from IRB review, although many IRBs insist on reviewing exempt research protocols to confirm that they are exempt. The exemptions are quite limited, and several impose restrictive confidentiality requirements. For example, one provision broadly exempts "survey procedures, interview procedures, or observation of public behavior," unless

- information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and
- any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

IRBs have interpreted this limitation quite expansively, meaning that the exemption has been quite narrow in practice.

Finally, although the IRB regulations technically apply only to research funded by the federal government, an institution that receives such funds must provide assurance that it will protect the rights and welfare of the human subjects of all its research on human subjects, whatever the source of funding. This mismatch is non-trivial; nearly 80 percent of all research projects reviewed by the University of Chicago's Social Science IRB are either personally funded, privately funded, or unfunded. Many academic institutions have declined to commit themselves to using the Common Rule framework for reviewing non-federally funded research, but nonetheless seem to have adopted the IRB framework to review such research.

FALSE POSITIVES AND FALSE NEGATIVES

IRBs exist to perform research oversight. Any system of research oversight will generate four kinds of results: true positives, true negatives, false positives, and false negatives. True positives and true negatives occur, respectively, when research that should have been approved is approved, and research that should not have been approved is not approved. False positives and false negatives occur, again respectively, when research that should not have been approved is approved, and when research that should have been approved is not approved. True positives and true negatives are correct results; false positives and false negatives are mistakes. (In scientific circles, a false positive is called a Type I error; a false negative is called a Type II error).

The goal for any system of research oversight is to maximize the number of true positives and negatives, and minimize the number of false positives and false negatives and the costs of research oversight. Those costs include the transaction costs of operating the system (e.g., researcher time spent filling out forms and answering questions, the salaries for staff hired to deal with IRB compliance, the salaries of IRB personnel, the time that is volunteered to review protocols) and the costs of erroneous decisions and delay.

No system of research oversight will ever operate perfectly, but different strategies impose different costs. A commitment to screen every protocol exhaustively lowers the frequency of false positives, but increases the frequency of false negatives — and delays research across the board. Conversely, a commitment to expeditious approval lowers the frequency of false negatives and minimizes delay, but increases the frequency of false positives.

ASYMMETRY In principle, there is no reason to prefer false negatives to false positives, or vice-versa, as long as the harms are symmetrical. From this perspective, the goal is to minimize the joint sum of administrative and error costs, irrespective of the source of those costs. In practice, however, false positives are viewed as a much more serious problem than false negatives. False positives that result in adverse outcomes are highly salient — and hindsight bias encourages everyone involved to view (*ex ante*) close cases as (*ex post*) clear mistakes. Because false positives can result in lawsuits, adverse publicity, government action, and disciplinary proceedings, they are self-revealing examples of institutional failures.

False negatives are not subject to the same dynamic. The victims of a false negative decision are the researcher and those who might have been helped by the research. The researcher has little incentive to pick a fight with the IRB or draw public attention to the dispute because there is a significant opportunity cost in doing so — particularly when the researcher will have to get approval from the IRB for his or her next project. Those who might have been helped by the research are a diffuse group of invisible victims who may not even know about the research. Even when there is a vocal group of identifiable victims, there is no one to whom to appeal an adverse IRB decision.

This asymmetry means that institutions have a strong incentive to drive the false positive rate down, without weighing the cost of the increased false negative rate and the delays and transaction costs of review. This dynamic is compounded by the scandal-mongering that dogs these controversies. A high-profile false positive decision will cause university administrators to put pressure on their IRBs to avoid such problems — and IRBs will respond by becoming even more gun-shy than they are to begin with. These problems are compounded by mismatched

incentives and the externalization of many of the costs of research review. The result is that there is no feedback loop with regard to the rate of false negatives and the costs of compliance — so both tend to increase without effective constraint. Stated more concretely, there is a one-way ratchet in favor of fewer false positives and more false negatives.

The resulting dynamic is nicely described in Henry I. Miller's book *To America's Health.* Miller, a former high-ranking official in the U.S. Food and Drug Administration, writes that a false positive

> *mistake is highly visible and has immediate consequences — the media pounces, the public denounces, and Congress pronounces. Both the developers of the product and the regulators who allowed it to be marketed are excoriated and punished in modern-day pillories: congressional hearings, television news magazines, and newspaper editorials. Because a regulatory official's career might be damaged irreparably by his good faith but mistaken approval of a high-profile product, decisions are often made defensively — in other words, to avoid [false positive] errors at any cost.*

The results are both predictable and perverse, writes Miller:

> *In the early 1980s, when I headed the team at the FDA that was reviewing the NDA [new drug application] for recombinant human insulin,... we were ready to recommend approval a mere four months after the application was submitted (at a time when the average time for NDA review was more than two and a half years). With quintessential bureaucratic reasoning, my supervisor refused to sign off on the approval — even though he agreed that the data provided compelling evidence of the drug's safety and effectiveness. "If anything goes wrong," he argued, "think how bad it will look that we approved the drug so quickly."*

Stated more concretely, a false positive is costly to the IRB, but a false negative is costly to the individual researchers and those who would have benefited from the research. An IRB will accordingly have an incentive to drive down its false positive rate and be (relatively) indifferent to its false negative rate, while researchers will tend to have the opposite incentives.

If those problems sound familiar, it is because they are. Decisionmakers always confront a tradeoff between false positives and false negatives when considering whether to approve a drug, establish the criteria for obesity, use a per se test for assessing whether conduct violates the antitrust laws, prohibit airplane passengers from bringing liquids aboard, declassify (or classify) government documents, prosecute a person (and for what), set a specific speed limit, establish a level of blood alcohol that constitutes drunk driving, and so on.

IRB PERFORMANCE

Too often, policy debates on the "best" decision rule for handling such problems rely on platitudes and anecdotes. Such rhetoric provides no help in determining the optimal tradeoff between Type I and Type II errors for any given problem. Careful assessment of the costs and benefits of different decision rules is much more likely to result in good public policy than any of the alternatives. Accordingly, I now turn to what we know about IRB performance.

BENEFITS There is no empirical evidence that IRBs have any benefit whatsoever.

COSTS There is a modest literature on direct IRB costs. The most comprehensive study (from 2005) looked at the costs incurred by 63 academic medical centers with IRBs. It found that annual operating costs ranged from $171,000 to $4.7 million, with a median cost of $742,000.

The cost per reviewed protocol ranged from $431 to $644. Expedited reviews were as expensive as full protocol review — most likely because of the high opportunity cost of having senior IRB personnel perform the expedited review. Earlier studies found smaller costs per reviewed protocol — most likely because they did not account for the opportunity cost of the time of those who volunteer to serve on an IRB.

Several studies have examined the costs of obtaining IRB approval in multi-center studies. One analysis involved an observational study of 43 U.S. Department of Veterans Affairs primary care clinics. Approximately 4,700 hours of staff time over a 19-month period were devoted solely to the IRB process, with the time required to obtain approval varying from 52 to 798 days, with a median of 268 days. The second study concluded that 17 percent of the total research budget in a multicenter study was spent obtaining approval from seven additional IRBs after the "home" IRB had approved the protocol — a process that took 18 months to complete.

OUTPUT Numerous studies have documented considerable variability, inefficiency, and uneven protection of study participants in multi-center studies, where the same protocol is reviewed by multiple IRBs.

In one such study, different IRBs used different standards to review the research protocol. Although the study was designed to qualify for expedited review, one site exempted it from review, 31 required full review, and one rejected it outright. Twelve sites requested (and two insisted on) provisions that increased the risks to study participants. The IRBs were fixated on paperwork and required researchers to submit multiple revisions of applications, consent documents, and ancillary forms — even though most revisions involved minor, non-substantive editorial changes to the wording of the consent document. The authors dryly concluded that "several features of the IRB system as currently configured impose costly burdens of administrative activity and delay... and paradoxically decrease protection of human subjects."

In another study, the IRBs demanded many changes in the formatting and wording of the consent and survey forms — and each change demanded by one IRB had to be approved by all the others. The researchers in this study asserted that by the end of the process, no substantial change had been made in the protocol, and that the changes demanded had no discernible impact on the protection of human subjects. Instead, there were "changes as trivial as saying 'study description' rather than 'description of study,'. . . [and they] 'spent thousands of dollars doing things like changing the font from Helvetica to Times New Roman and making sure the border of the forms were the right color.'" Other researchers

have similarly noted that IRBs seem to spend most of their time tinkering with consent forms — generally making them longer, less readable, and introducing errors and delay. Similar criticisms have been made by the Institute of Medicine and the National Bioethics Advisory Commission.

These problems are not limited to multi-center studies. A survey of IRB chairpersons also found wide variation in the application of risk and benefit categories for pediatric research, with some of the determinations at odds with the IRB enabling regulations — and with the available data on risks.

Finally, one study submitted three flawed biomedical research protocols to 32 IRBs at major universities. Twenty-two IRBs participated in the study. The study found substantial inconsistency "in the application of ethical, methodological, and informed consent standards." The authors noted that "our evidence supports the conclusion that IRBs approve inappropriate investigations and inhibit appropriate investigations. Deficient protocols may be modified, but only the most glaring problems are identified." A separate description of the study by one of the authors was more cutting: "We submitted three flawed medical protocols with the expectations that the review boards would find the flaws evident in the proposals. But the boards, on the whole, did not find the flaws. Rather, they suggested changes that, if approved, [would have] made bad research projects worse."

To summarize, there is a fairly substantial body of research indicating that IRBs operate inconsistently and inefficiently, and focus their attention on the submitted paperwork and consent forms.

COMPLAINTS AND RISKS There are a host of more generic complaints about IRBs, including delay, increased costs, and useless paperwork. There have also been complaints about IRBs discouraging research in particular areas and harassing individual researchers. Sometimes this process is triggered by pressure from third parties, while other times it is instigated by university administrators or IRB members.

Scholars in the social sciences and humanities have been particularly vocal about their concern that the IRB model, which was designed to review biomedical research, does not "fit" their research methodologies. Law professors have started to pay attention to these issues as well. There is no question that nonsensical results can follow from the unfortunate combination of good intentions, risk-averse lawyers, and bungling bureaucrats. Such incidents are ripe for caricature, but they help inform perceptions about the value added by IRBs — or lack thereof. Academics who practice oral history, ethnography, and journalism have been particularly vehement about the risks IRBs pose to their respective disciplines.

WORKLOAD Concerns have been expressed about the adequacy of IRBs' funding and heavy workload. The largest IRBs review thousands of protocols per year, along with hundreds of reports of adverse events. At Johns Hopkins University, until June of 2001, a single IRB, meeting once every two weeks, was responsible for the approval of 800 new protocols annually and the continuing monitoring they generated.

Some of the workload problems appear to be the result of "mission creep," as IRBs assume (or are given) responsibility for more things. Consider the fact that many IRBs review research that is exempt from IRB review to determine whether the research is, in fact, exempt. At many institutions, IRBs have also been given responsibility for assessing compliance with federal privacy laws. As noted above, IRB review at each site of a multi-center study results in an obvious duplication of effort and increase in the aggregate IRB workload.

SUMMARY The available evidence indicates that there are substantial direct and indirect costs associated with IRB oversight of research. IRBs also operate inconsistently and inefficiently, and focus their attention on paperwork and bureaucratic compliance. Despite their prevalence, there is no empirical evidence IRB oversight has any benefit whatsoever — let alone benefit that exceeds the cost.

SUPPLY AND DEMAND

How did we end up with thousands of IRBs, each busily reviewing protocols and giving rise to considerable dissatisfaction among those they regulate? Understanding the development of the current state of affairs requires separate examination of the supply and demand sides of the dynamic.

On the supply side, the spread of IRBs is a typical example of how bureaucracies behave. The Common Rule represents the attempt of one bureaucracy (agencies within the federal government) to influence the behavior of another bureaucracy (academic medical centers and universities). The federal government leveraged the receipt of federal dollars into a requirement that these institutions create a new bureaucracy (IRBs) that would perform research oversight.

Once IRBs were created, they channeled and regularized the mandated demand for their services by requiring researchers to fill out forms that the IRBs could then review. The use of forms as the locus for decisionmaking resulted in a "check-the-box" mentality among IRB personnel, where the substantive content of the research was less important than the forms that had been submitted. The fixation on forms also encouraged IRB personnel to become obsessed with the language and formatting of consent forms. Those tendencies are worsened by the fact that evaluation of IRB performance is based on whether the forms and other paperwork were filled out and IRB decisions were properly documented.

Consider what happened after the death of a young and healthy research subject at Johns Hopkins University. The review by the two federal agencies with responsibility for IRB oversight focused on the failure of the IRB to maintain adequate minutes and the fact that Hopkins' reliance on executive subcommittees meant that "most protocols are neither individually presented nor discussed at a convened meeting of any IRB." From a bureaucratic perspective, this approach makes perfect sense. Matters of form are readily verifiable and objective, while the actual substance of IRB review is not. The nexus of these matters with the actual issue at stake (the death of an otherwise healthy research subject) was quite another matter. No time was spent considering whether better minutes and less use of executive subcommittees would have made the

slightest difference in the tragic outcome. However, this emphasis on things that could be easily measured and reported (form completion, minutes, and use of subcommittees), instead of things that are less easily measured and reported (whether the IRB appropriately balanced the risk of Type I and Type II errors) is typical of bureaucracies.

Bureaucracies also tend to expand to capture all available resources and regulatory space. This tendency explains why IRBs ended up being the focal point for research oversight of non–federally funded research, and many IRBs insist on reviewing exempt research to determine whether it is exempt. Thus, the creation of a system for research oversight created a demand for more oversight — with associated mission creep.

The scandal-based origins of the Common Rule also affected the nature and direction of the enabling regulations. The Common Rule is exclusively concerned with the protection of human subjects, and gives no weight whatsoever to the academic freedom of researchers or to the social costs from research that is delayed, constrained, or not performed at all.

Because bureaucracies are risk averse, the threat that someone somewhere might sue has encouraged IRBs to lawyerize the research oversight process. This dynamic furthered the fixation of IRBs on increasingly detailed consent forms and the protection of study participants against increasingly remote and trivial risks.

IRBs also tried to avoid other types of controversy by adopting politically correct attitudes toward proposed research. One study of the IRB review process found that high-quality studies involving "sensitive" topics were rejected as often as flawed studies of non-sensitive topics — with the "potential political impact of the proposed findings" given as the primary reason for rejection of the high-quality studies. For example, one IRB rejected a study of reverse discrimination on the grounds "the findings could set affirmative action back 20 years if it came out that women with weaker vitae were asked to interview more often for managerial positions than men with stronger vitae."

Finally, the structure of IRB oversight means that each institution's IRB has an effective monopoly on research oversight at that institution. The result is that there is no competitive constraint on IRB inefficiency, incompetence, political correctness, and other forms of opportunistic behavior.

On the demand side, there were few natural constraints on the expansion of IRBs because researchers have little incentive to make a fuss — particularly when doing so marks one as a troublemaker the next time through the process. Time spent disputing with the IRB is also time not spent on research.

These supply- and demand-side considerations reinforce one another, and result in the dysfunctional state of affairs that we find ourselves in today.

JUDICIAL OVERSIGHT OF IRBS

As noted previously, multiple lawsuits have been filed in recent years against IRBs. What happens when judges are invited to second-guess the research oversight process? Most of the lawsuits have not resulted in legal opinions, so it is hard to know how judicial oversight will ultimately play out — but if one high-profile example of such litigation is any indication, judicial self-righteousness (if not out-and-out incompetence) will create problems in handling such cases.

In *Grimes v. Kennedy Krieger*, the highest state court in Maryland vented its collective spleen on a well-designed and sophisticated study intended to assess the comparative cost-efficacy of various strategies for remediation of lead paint. The problem was an important one: 95 percent of the housing stock in the inner city neighborhood where the study was conducted had high levels of lead dust that had not been abated. In many instances, the cost of full remediation exceeded the value of the housing, so any attempt to compel such remediation would result in the abandonment of the housing, making everyone worse off. Conversely, a cost-effective strategy for remediation had the potential to make everyone better off — a fact that helps explain why the study was funded by two federal agencies (the Department of Housing and Urban Development and the Environmental Protection Agency) and administered in cooperation with the Maryland State Department of Housing and Community Development. (Indeed, the project has been replicated by HUD in 2,600 houses in 13 other cities.) The study was approved by the EPA (after internal and external reviews), the Centers for

Stupid IRB Tricks

- A linguist seeking to study language development in a pre-literate tribe was instructed by the IRB to have the subjects read and sign a consent form before the study could proceed.
- A political scientist who had bought a list of appropriate names for a survey of voting behavior was required by the IRB to get written informed consent from the subjects before mailing them the survey.
- A Caucasian Ph.D. student, seeking to study career expectations in relation to ethnicity, was told by the IRB that African-American doctoral students could not be interviewed because it might be traumatic for them to be interviewed by the student.
- An experimental economist seeking to do a study of betting choices in college seniors was held up for many months while the IRB considered and reconsidered the risks inherent in the study.
- An IRB attempted to block publication of an English professor's essay that drew on anecdotal information provided by students about their personal experiences with violence because the students, though not identified by name in the essay, might be distressed by reading the essay.
- A campus IRB attempted to deny a master's student her degree because she did not obtain IRB approval for calling newspaper executives to ask for copies of printed material generally available to the public.

SOURCE: *AAUP Thomson report*

Disease Control, and the Johns Hopkins University IRB. The IRBs for 29 different entities in each of the 13 other cities had also approved the protocol.

Although the degree of remediation varied (which was, after all, the point of the study), all study participants lived in a remediated apartment. Individuals could live in the remediated apartments without participating in the study, and they were solicited for participation in the study only after they were already living in a remediated apartment. Even if an individual did not participate in the study, he or she got to live in a remediated apartment that had a better lead paint environment than the alternatives that were available to 95 percent of the affected population.

The lawsuit was filed by several research subjects, alleging that they had been exposed to unsafe levels of lead because of their participation in the study. The trial court granted summary judgment to defendants on the grounds that they did not have a generalized duty of care to study participants. The only issue before the appellate court was whether or not there was a generalized duty of care. There was a sparse factual record, and the larger context of the dispute involved the balancing of complex issues of public policy. An amicus brief filed by a third party canvassed the history of research abuses and argued that the court should find that there was a duty of care and send the case back for trial.

Rather than address the narrow issue before it, the Court held that the research was per se inappropriate, unethical, and illegal. Relying on suspicion, innuendo, and a clear misunderstanding of the facts and the law, the Court smeared a well-respected researcher and the IRB with its flat assertion that the study "differs in large degree from, but presents similar problems" as, the Tuskegee Syphilis Study and the "research" performed in Nazi concentration camps. The Court also felt compelled to remind the researchers that children were not "rats, hamsters, monkeys, and the like" — even though there was no evidence the researchers thought otherwise.

The opinion also included language that would have closed down all non-therapeutic pediatric research in Maryland — notwithstanding a federal statutory commitment to the contrary, and the fact that such research accounted for millions of dollars of funding to Maryland institutions. A motion for reconsideration pointed out these problems, using the understated opening line that "on the day the mandate in this case issues, hundreds of fully accredited medical research projects now conducted in Maryland will terminate." A flurry of amicus briefs by prestigious organizations made the same point — including one from the third party that had started the whole process, but that had come to believe the Court had gone completely off the deep end. In a press release, Johns Hopkins University noted the effect of the ruling would be "the loss of valued researchers and investigators, who will be forced to relocate elsewhere in order to conduct their research."

In response to this outpouring of criticism, the Court denied the motion for reconsideration, but unpersuasively argued that critics had simply misunderstood the original opinion. At no point did the Court respond to the concurrence (issued with the original opinion), which made it crystal clear that the original opinion said what it meant and meant what it said. After the case was returned to the trial court, it was ultimately dismissed with prejudice — an unimpressive (if not ignominious) ending for a case where the state supreme court condemned the conduct of the defendants by analogizing it to Nazi research.

Anyone who thinks the quality of research oversight by IRBs will necessarily be improved by judicial oversight should read the opinion. Strikingly, the Court's analysis was so over the top that it was harshly criticized by the assistant attorney general in the Maryland attorney general's office responsible for health policy, who stated that "the Court's rhetoric was heated, its historical comparisons inflammatory and unjust, and aspects of its decision ill-considered." He also noted that the Court's opinion was marked by "rhetorical excess and [a] condemnatory tone... [and it] suffers badly from imprecision and superficiality." With friends like that, who needs enemies?

A COMPARATIVE INSTITUTIONAL PERSPECTIVE ON IRBS

All human institutions operate imperfectly. Expecting IRBs to prevent all bad outcomes or perform flawlessly is foolish. Indeed, a non-zero incidence of "bad outcomes" is fully consistent with an optimal level of research risk and research oversight. As the dean of the Johns Hopkins School of Medicine noted, the alternative "is not to do any clinical investigation... and still have children on ventilators after polio."

The question we should be asking is whether IRBs constitute the "least worst" institutional response to the problem of balancing the marginal cost and marginal benefit of research and research oversight. As the foregoing suggests, in several important respects the answer to that question is almost certainly "no."

This assessment should not lead to regulatory nihilism. Even if the Common Rule had never been enacted, there are good reasons to think that most research institutions would have developed mechanisms for reviewing the work done by their personnel. However, it is unlikely that each institution would have arrived at the solutions dictated by the Common Rule — let alone the obsession of the Common Rule with the race, gender, cultural background, and affiliation of those serving on the IRB. Instead, it seems probable that a diverse array of forms of research oversight would have emerged, and that they would have:

- varied significantly across institutions, depending on local culture and the degree of risk-aversion;
- varied significantly across disciplines and research modes;
- treated different researchers differently, based on the relative risk of the underlying research and the reputation and experience of the individual researcher; and
- varied in the relative amount of *ex ante* review and *ex post* enforcement.

As this short list suggests, there is no compelling reason why universities and academic medical centers should be shackled to the Common Rule in evaluating non–federally funded research — particularly given the potential gains from greater diversity in the forms and modes of research oversight.

IMPROVEMENT The balance of this section accordingly offers some suggestions for improving the performance of research oversight.

Make the invisible visible. We have some empirical evidence on the cost of operating an IRB, modest empirical evidence on the cost of obtaining IRB approval, and no empirical evidence whatsoever on the social cost of false negatives and the social benefits of IRB oversight. Any attempt to assess the value of IRB oversight requires quantification of the costs and benefits. It is long past time for IRBs to be an object of study, instead of simply the tollgate through which all studies must pass.

Make the cheap expensive. Most of the cost of research oversight is externalized. The federal government requires the use of IRBs for federally funded research, but it does not pay its fair share of IRB costs. The federal government also strongly encourages the use of IRBs for non–federally funded research, but it pays nothing toward those costs. If the federal government wants IRB review, it should pay for it.

At the level of individual institutions, the reliance on volunteers means that even the direct costs of IRB review are not fully captured in the institution's budgeting process. Similarly, an institution's budgetary expenses for IRBs do not reflect the social costs of delay and false negative decisions. If individual institutions want to rely on IRB review, they should face up to its full cost. Stated differently, only by making the cheap expensive (i.e., by making off-budget expenses on-budget) will we be able to assess the true cost of the IRB review process.

Make the expensive cheap. Although certain research methodologies are exempt from IRB review, there are significant constraints on the scope of those exemptions — and many IRBs require submission of exempt research protocols in order to determine whether they are, in fact, exempt. Instead of this approach, as a recent American Association of University Professors (AAUP) report suggests, "research on autonomous adults whose methodology consists entirely in collecting data by surveys, conducting interviews, or observing behavior in public places should be exempt from the requirement of IRB review — straightforwardly exempt, with no provisos, and no requirement of IRB approval of the exemption." This approach will save researchers a great deal of time and effort. It will also eliminate a considerable amount of totally unnecessary work currently done by IRBs — freeing them to focus on projects that impose greater risks.

The existing framework for exemptions purports to ensure more confidentiality than would otherwise be the case — but even that "benefit" is likely illusory. IRB oversight will only add value to existing departmental and disciplinary practices for ensuring confidentiality if IRB members are better equipped to assess practices for collecting and storing data than members of the discipline are — a proposition that is dubious on its face. Worse still, IRBs have expansively interpreted the requirement that exemption is not available if "any disclosure of the human subjects' responses outside the research could . . . be damaging to the subjects' financial standing, employability, or reputation." As the AAUP report noted, it is hard to see a clear line between cases in which a breach of confidentiality might be damaging to the subjects' financial standing, employability, or reputation, and cases in which it would not be. It is also hard to see why full-blown IRB review is necessary for surveys and interviews when autonomous individuals can decide whether or not to participate — and there is no obvious justification for full-blown IRB review of cases involving observation of public behavior.

Stick to thy last. The point of the IRB framework was to protect research subjects — not to engage in censorship at the behest of vocal interest groups or the ideological preferences of IRB members or campus administrators. The Common Rule flatly prohibits such stratagems: "[T]he 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." However, this clear prohibition has not stopped IRBs from doing precisely the opposite.

If an institution does not want its personnel to perform certain research because it is unprepared for the political heat, it should forthrightly announce that it does not want to be associated with such projects. The institution should not hide behind the purported desire to protect human subjects — and IRBs should stop enabling such conduct.

Destroy local monopolies. The current regulatory framework effectively creates institution-specific monopolies for each IRB. Such monopolies are inherently undesirable. A robust market in IRB review would create better incentives than the current system, and would keep the review process and reviewers in line. It would also provide a market test of the claim routinely heard from IRB members and their supporters that IRB oversight improves the quality of the research that is conducted. Any IRB meeting the requirements of the Common Rule should be authorized to perform reviews of research protocols, irrespective of where the research will be performed. The Common Rule authorizes a version of this proposal for multi-center studies, so this proposal simply extends this model to single-center studies. As in the current system, institutions would be free to scrutinize the IRB decision and turn down the proposed research — but they could not do so on the grounds that it was not approved by their IRB.

Just say no to bioethics. IRBs are one of the few success stories of bioethics — an institutionalized arrangement for the evaluation of the ethical implications of all research conducted by every entity that receives federal funding. There may be a bioethics cop on the beat, but as noted previously, it has proven itself prone to a host of predictable pathologies. IRBs also provide a rhetorical hook for criticism and an institutional focus for liability; instead of criticizing an individual researcher, the fact of IRB approval means that the institution as a whole is on the line for any shortcomings. IRB approval provides no protection against criticism by novice bioethicists looking to "make their bones" on the latest controversy or

established bioethicists who want to see their names in print. For example, the Kennedy Krieger study was criticized by several bioethicists on the grounds that the researchers should not have accepted the proposition that resource constraints made full abatement unlikely. From that exalted perspective, trying to find more cost-effective abatement strategies meant that the researchers were, in fact, capitulating or colluding with "status quo conditions of gross social injustice." How could anyone believe that "sticking it to the man" was a better strategy than actually doing something to help the children of Baltimore? As George Orwell has noted, "[O]ne has to belong to the intelligentsia to believe things like that: no ordinary man could be such a fool."

Wholly apart from these problems, it is hard to make the case that IRBs, with their obsession with paperwork and the tweaking of consent forms, actually promote the protection of human subjects. The next time a bioethicist comes calling with a bright idea, just say no.

CONCLUSION

As noted previously, the right question to ask about IRBs is not whether they are perfect, but whether they are the "least worst" institutional response to the problem of balancing the marginal cost and marginal benefit of research and research oversight. Even judged by this modest standard, IRBs fall well short. Minimal improvements have the potential to move us closer to a "least worst" world and constrain the excesses that currently prevail in research oversight.

What role is there for lawyers, law professors, and judges in this process? There is a general tendency in legal circles to believe that "things go better with due process," and the keys to the courthouse are the keys to the kingdom. Those inclined to such views should be required to write 100 times on their respective blackboards Professor Grant Gilmore's warning against such tendencies: "In heaven there will be no law and the lion will lie down with the lamb.... In hell, there will be nothing but law, and due process will be meticulously observed." R

Readings

- "Don't Talk to the Humans," by Christopher Shea. *Lingua Franca*, Vol. 10 (2000).
- "Improving Protection for Research Subjects," by Robert Steinbrook. *New England Journal of Medicine*, Vol. 346 (2002).
- "Improving the System for Protecting Human Subjects: Counteracting IRB 'Mission Creep,'" Center for Advanced Study, University of Illinois College of Law, 2005.
- "Protecting Research Subjects: The Crisis at Johns Hopkins," by Robert Steinbrook. *New England Journal of Medicine*, Vol. 346 (2002).
- "Research on Human Subjects: Academic Freedom and the Institutional Review Board," by Judith Jarvis Thomson et al. *Academe*, Vol. 91, No. 5 (Sept.–Oct. 2006).
- "U.S. Health Researchers Review Their Ethics Review Boards: A Qualitative Study," by Scott Burris and Kathryn Moss. *The Journal of Empirical Research on Human Research Ethics*, Vol. 1 (2006).