CHAPTER 3

Weighing Risks and Benefits, and Undue Inducement

Weighing risks and benefits can seem straightforward, but is at times extremely complex.

In June 2013, one of the most divisive recent debates in bioethics erupted concerning the risks and benefits of a study that examined how much oxygen to give premature infants. Every year, approximately 500,000 babies are born prematurely, and 5,000 of them die.[1,2] The vast majority of premature and low-birth-weight infants require supplemental oxygen.[3] Pediatricians in Neonatal Intensive Care Units (NICUs) have, however, faced a dilemma. If they give too much oxygen, babies are more likely to survive, but also to develop vision problems and become blind. If the doctors give too little oxygen, more babies die, but the survivors are more likely to be able to see. Thus, doctors give a wide range of oxygen levels, hoping to avoid either too much or too little. But within this wide spectrum, doctors vary. Some choose 89 percent oxygen saturation, while others give closer to 95 percent. Until 2004, no study had been conducted to determine definitely which was, in fact, better.

Hence, from 2004 to 2009, 23 institutions, including Yale, Stanford, and Duke, conducted the Surfactant, Positive Pressure, and Oxygenation Randomized Trial (SUPPORT) study. IRBs at all 23 institutions approved the study. When the mothers were in labor, researchers randomly assigned the future infants, before they were born, to one of two levels of oxygen saturation: 85–89 percent or 91–95 percent. In May 2010, the results were published,[4] showing that the higher amount was more likely to avoid death, but also to cause blindness. At the higher level, 16.2 percent died before leaving the hospital and 17.9 percent developed eye disease. At the lower level, more died (19.9 percent), but fewer (8.5 percent) developed eye disease.

Two years later, in 2011, the Office for Human Research Protections (OHRP) received a complaint about the study and, in response, sent a letter to the investigators. The agency wrote that the informed consent forms had been inadequate, and asked the institutions to take corrective action.[5] Public Citizen, the consumer advocacy group founded by Ralph Nader, vociferously attacked the study and argued that OHRP had "failed to demand adequate corrective actions."[6] A *New York Times* editorial, entitled "An Ethical Breakdown," concluded that the study's failure was "startling and deplorable."[7] A major battle erupted, with competing groups of 45 and 46 bioethicists each publishing opposing letters to the *New England Journal of Medicine* in June and July 2013. In December 2013, *The American Journal of Bioethics*

devoted an entire issue to the controversy, publishing three major articles, and 17 responses from leading bioethicists—disagreeing significantly on which side was right and why.[8]

OHRP argued that the study's informed consent forms significantly overemphasized the benefits and downplayed the risks of the study, and failed to mention death as a possible outcome. The forms said, for example, that "subjects may have a possible decrease in chronic lung disease…and/or a decrease in the need for eye surgery"[9] But these documents did not mention the possibility that the subjects, conversely, may have an increased risk of lung disease or eye problems. OHRP also stated that the study was not minimal risk, as the researchers had contended, but rather involved substantial risks since many infants would receive oxygen levels that were different than those they would have otherwise gotten.

Some observers supported OHRP, concurring that the consent forms had significant deficiencies. But the study's defenders maintained that, in questioning this study, OHRP had "over-reached."[10] They argued that death was a possibility for these premature infants even if they were not in the study; and that these consent forms only needed to mention those aspects of the study that were *not* part of the care that participants would receive anyway.[11] Since the risk of death for these infants was not greater than it would have been outside the research, they said that the consent form did not need to mention this danger. Moreover, these study advocates averred that when doctors are divided about the use of two different treatments and both approaches are considered "standard of care," studies are needed to decide which is better. Such comparative effectiveness research (CER), as it is called, is vital to improving health care. These proponents argued that this study exemplified such research, and that to oppose it would impede further such investigations. In fact, Simon Whitney, a physician and lawyer at Baylor Medical School, had argued that for this study, informed consent could be waived altogether, since the research was minimal risk.[12] He pointed out that obtaining consent for this project cost $200,000, and took the study more than twice as long to enroll the needed number of subjects. In the meantime, without the results of the experiment, many infants would continue to die. Since doctors were divided, both levels of oxygen were acceptable. Randomizing infants to one of these two doses may change which one each patient would otherwise receive but, the study's supporters argued, does not in itself necessarily make that treatment better or worse.

Opponents countered that randomization itself *was* a risk. In the study, modified oxygen machines controlled and masked gas levels in ways that do not usually occur. Hence, physicians would not be able to monitor and adjust the infants' level over time as they otherwise would. Randomization, rather than the clinician's judgment about each individual patient, would determine the dose, which would be fixed through the course of the treatment.

In addition, a central tenet of research ethics is that subjects can be randomized to one of two interventions only if so-called clinical equipoise exists—that is, if scientists do not have reason to believe beforehand that one approach is better than the other. If the researchers think one treatment is superior, they should not knowingly be giving subjects the other, inferior treatment. Yet earlier research, on which the SUPPORT study was based, showed that 93 percent of NICUs used higher, rather than lower doses. Only 7 percent of NICUs had maximum targets of less than 92 percent, and 93 percent had targets of more than 92 percent. Moreover,

some hospitals use a much narrower range of 88–92 percent. Hence, in assigning many infants to lower amounts of oxygen (85–89 percent), the study increased their risk of death.[13] Critics have thus argued that equipoise did *not* exist. Presumably, many parents would not want their premature baby randomized to receive a level that only 7 percent of centers provided (while 93 percent of centers gave more), knowing that the risk of death would probably increase. Granted, lower doses would presumably reduce the risk of blindness. But to many parents, these two risks would undoubtedly not be equivalent—they would understandably prefer to have a blind child than a dead one.

Ideally, if doctors are divided in their practices 50/50, randomizing patients to each treatment is fine; but if 93 percent of doctors prescribe a higher level, and only 7 percent aim at a lower amount, questions emerge about giving 50 percent of the ill subjects the less popular dose. The consent form stated that both levels were considered "acceptable." The parents were *not* told that 93 percent of doctors give the higher oxygen level. Half of the parents might not want their child to have only the lower level, if they knew that only 7 percent of doctors provided it. Critics argued that the study, as it was designed, did not have equipoise, and hence should not even have been conducted.

Debates continue as to whether randomization was a risk. John Lantos[14] argued that in earlier studies of oxygen saturation, infants in studies did better, regardless of which treatment they received, than did those outside the study. Nurses and doctors may be extra-motivated and devote special attention to patients in a study. Therefore, Lantos wrote, randomization should in fact be listed as a benefit of studies.[xiv] But as Public Citizen points out, infants in SUPPORT were healthier than those excluded by researchers from the study,[15] and hence did better—not because study subjects received extra attention in the study. Moreover, in many other studies, patients receiving a new treatment do *worse* than those receiving standard of care.

These debates thus open larger questions of when exactly equipoise exists—exactly *how* split doctors need to be about two competing treatments to conclude that *patients can be randomly assigned to each*. In these debates, each side seems to raise several valid points, but neither group is entirely right or wrong. Though study proponents argue that critiques of SUPPORT imperil all comparative effectiveness research, the question is not whether such research can be conducted, but *how*—exactly which treatments should be compared, and what subjects should be told about them.

At the same time, though OHRP asked the University of Alabama to take corrective action, it is not clear what action would be appropriate at this point.

In June 2013, however, under intense pressure from NIH and others, OHRP backed down. These debates continue, and parents whose infants died are suing the researchers, bolstered by OHRP's charge of unethical behavior

As I saw when my father wrestled with whether to undergo chemotherapy with its terrible side effects and unclear benefits, weighing potential unknown risks and benefits can be extremely hard—especially when life and death are involved. Even in one family, individuals can differ widely about participating in research. IRBs, I soon found, often struggle and argue about these questions, too.

Since the federal regulations governing IRBs are relatively minimalist, individual

committees have a great deal of leeway concerning these issues. Law, like ethics, is not a science, but is bound by many more rules and seeks to follow precedents very seriously. Attorneys and even judges can interpret laws in different ways, but appellate courts exist with the sole purpose of adjudicating and often overturn the rulings of lower judges. The Supreme Court serves as a final arbiter, though its nine justices are divided rather than unanimous in 70 percent of cases.[16] Moreover, even when the judges are unanimous, they may disagree widely on the legal reasoning behind their conclusions.[17] Interpreting and applying ambiguous language, these courts draw on published decisions, and document and disseminate their own interpretations, thereby establishing and building on widely published precedents. Judges regularly refer to this large body of case law. But for IRBs, no such mechanisms exist. Committees and researchers lack these tools. *Instead, in many ways, local IRBs serve as their own police, judge, jury and Supreme Court*. Institutions and OHRP can affect IRBs. Nevertheless, these committees maintain enormous autonomy.

Once a committee has determined how to interpret and apply the regulations for a particular study, it tends to see its decision as incontrovertible, rather than as subject to differing alternative interpretations. Even many philosophers have seen principles as either present or absent in an argument. While some philosophers feel that for any ethical problem or question, only one best answer exists, in practice wide disagreements persist about many areas in bioethics, even among experts. Every month, *The American Journal of Bioethics* publishes target articles followed by six to 12 commentaries that usually offer a range of contrasting, if not dramatically conflicting, views.

Balancing Risks *vs*. Benefits

When evaluating a study, one of the IRB's chief charges is to examine and weigh the potential risks to the patient versus the potential benefits—to the patient and to the expansion of scientific knowledge and society at large. IRBs often struggle with this balance. In general, researchers should follow the principle of clinical equipoise—a patient should be entered into a study comparing two therapeutic approaches only if the researcher is genuinely uncertain at the outset whether one of these treatments is better than the other.[18] However, applying this principle can be tricky.[19,20]

In general, people view probabilities subjectively, using biases and so-called heuristics—simplified ways of conceptualizing complex competing odds. As the Nobel Prize–winning psychologist Daniel Kahneman has argued, if a wild boar is suddenly running toward us, we instantly gauge whether it is better to run left, right or backward—climb the tree or run to the cave[21]—without precisely calculating and comparing the risks and benefits of each possible alternative. So, too, we tend to rely on gut feelings in assessing risks. Studies have shown that, psychologically, most people overrate rare but traumatic events, weighing responses to possible losses more than responses to potential future gains[21]—overvaluing risks and undervaluing benefits. So, too, in making risky medical decisions, patients and doctors

frequently face uncertainties, and make and rely on rough assumptions.[22,23] Patients may also differ from providers in perceiving risks and benefits, based on their respective education and past experiences.[24,25,26,27] Perceptions of danger and risk can also involve subjective elements related to cultural rather than simply individual fears. In *The Immortal Life of Henrietta Lacks*, for example, Rebecca Skloot explains that many African American residents of Baltimore in the 1950s believed that members of their community were sometimes snatched off the streets at night and used in medical experiments,[28] reflecting histories of grossly unequal treatment during and after slavery.

How IRB chairs, members, and staff should respond to these complexities—how they are to weigh risks and benefits from the researcher's and patient's perspectives and also their own—remains unclear. Often IRBs must weigh many possible risks and benefits simultaneously. Committees must decide whether researchers should be allowed to test any two drugs against each other, if the products are ethically equivalent (i.e., that so-called clinical equipoise is present). A new drug might reduce a patient's symptoms more than the standard treatment, but cause worse side effects. Another drug might have fewer side effects, but be somewhat less effective. Yet the likelihoods of each of a study's possible future risks and benefits are frequently highly uncertain. The exact amounts of the possible differences—whether 10 percent more possible benefit is worth 20 percent more risk—can be crucial, but unpredictable and hard to weigh.

As we saw in the Kennedy Krieger study of lead levels in different homes (see Chapter 1), IRBs, courts, journalists, researchers, and subjects may perceive potential risks and benefits of a study differently. The court, and many journalists and observers, suggested that not only was the informed consent form deficient, but the study lacked equipoise—that the risks of only partially abated apartments were too great, and that the consent form inadequately described these.[29]

Yet, in making these assessments, individuals—whether IRB members, researchers, or subjects and their families—frequently rely on subjective "gut feelings" and highly personal assessments. As a nurse, Andrea was very aware of how idiosyncratically patients, families, and committees all perceive risks and benefits.

> It's very hard to weigh risks and benefits. …Everybody has to make calculations on their own. We thought a drug might help a sick patient, but a side effect was stroke. The likelihood was extremely small; but one patient turned it down because his mother had had a terrible stroke. He was a sick man—why would he turn down the possibility that this could help him for a 2 percent chance of a stroke? He wouldn't take that risk. All the IRB can do is try and make things as clear as we can. We have "likely," "less likely," and "rare but serious." I think "likely" is 20 percent, which in my mind is *not* likely. I would say "likely" is 50–60 percent—better than an even chance.

She felt strongly that IRBs need to ensure that consent forms conveyed these vagaries as best as possible. Yet *to describe risks not in percentages, but in descriptive language—as "likely" or "less likely"—is hard, and open to wide interpretation.*

DIFFERING DEFINITIONS AND THRESHOLDS OF RISK

"Truly Safe"? How Much Risk Is Okay?

When participants suffer or die in research studies, as did infants in the SUPPORT study, and Jessie Gelsinger in the gene transfer experiment, questions arise of whether these studies were too risky to have progressed as far as they had. IRBs therefore wrestle with exactly *how* safe a study is or should be. The difficulty of this question is compounded by the fact that the outcomes of experiments are by definition unknown, and risks can range from direct to indirect. Anxieties can thus shape IRB assessments of both the likelihood and seriousness of harm. Gauging potential benefits and harms is not a science, and is therefore shadowed by uncertainties. "When you don't know for sure what the risks and benefits will be, it's really hard, involving perceptions," Judy, a chair, explained. "You know what they *might* be, but you really don't know."

Consequently, I found, individuals and boards differ on the standard—whether a protocol needs to be *completely* safe to be approved, or whether the benefits only need to justify the risks. In fact, the regulations stipulate only the latter, not the former, and state that committees should reduce risks. Yet many IRBs feel they should "go beyond the regulations" and get rid of risks in some way. Olivia, the health care provider and chair, said,

> I worry: is the study *truly safe*? We do a lot of studies that are potentially high risk. We worry, but have to trust the investigator. That's why we look so carefully at the progress reports.

An IRB must rely on local researchers to do what they have claimed they are going to do, and to document what they have done. But establishing sufficient confidence can be hard. The IRB may not know of the researcher, or the PI may have a bad track record, or the risks can seem high. IRBs may be *overly* cautious, seeing their job as promoting subject protection, not science as well. Committees have ample reason to be vigilant, and no strong incentive to countervail this stance. Others feel that the cost of this vigilance to science is at times too high. Patrick, a physician and regular member, said,

> Being cautious is the IRB's job, but they may be *overly* cautious. For most IRBs, nothing good can come from approving a protocol. Every time you approve a protocol, there's a risk for bad things happening—including bad press.

Committees may thus avoid risk as much as possible, rather than lowering it to the point at which it is commensurate with the potential benefits.

IRB members and chairs who are not themselves investigators or clinicians may be especially wary of research. As an investigator who was new to his IRB, Patrick found this startling:

> Sometimes I get the impression that non-clinical IRB members and staff think that researchers are trying to harm people! Certainly, researchers have a vested interest and are biased—that's why it's good to have the IRB. I think HHS [The U.S. Department of Health and Human Services] says, "The protection of research subjects is the IRB's primary purpose." But that, almost by definition then, presents a tension in getting the research done.

In fact, Patrick misquotes the current regulations, which stipulate instead that risks should be weighed against benefits. But his misunderstanding is revealing; since he is a researcher, this impression comes probably from his IRB. He continued, "You hear this mantra of, 'What is the safest thing?' We're always going to err on the side of safety. But in clinical medicine, there's a justified risk/benefit ratio."

CHALLENGES IN QUANTIFYING RISK

Assessing dangers can be hard because many patients are already sick, like my father, and desperate for any interventions. Many ill patients also have what are called *therapeutic misconceptions*—even though they are participating in a blinded randomized study, they believe that their white-coated doctor is nonetheless choosing a treatment that is best for them individually, when in fact they may be randomized to a placebo. Despite being told otherwise, many patients persist in this belief.[30]

Participants may not understand the limitations and restrictions of research—especially in early-phase studies that are designed only to assess the side effects of drugs or initial possible effectiveness on small numbers of patients before being tried on larger numbers, and that usually have little promise of providing clear benefits. Andrea said,

> In their heart of hearts, subjects believe they're going to be the lucky guy who might get cured with this unlicensed therapy. Patients may agree to pretty onerous things because they're *desperate*.

Patients may fail to grasp the complexities involved, and just want to enroll. As Andrea said about consent forms,

> I'll say to subjects, "You should take it home and read it." "No, I don't want to read it. I just want to sign it." Some people are quite militant. They don't want to hear about it. They're scared of dying, and the doctor said he thought they might be interested in this. So they believe that the doctor is recommending what's best for them, even though the doctor may have clearly said, "I don't know if this is going to work." Patients don't necessarily hear what you tell them.

Patients with no other treatment alternatives may enroll in so-called last resort studies that have high risks and very uncertain potential benefits, posing additional challenges. Unfortunately, both experimental *and* existing therapies may be high risk and offer little benefit. IRBs must then struggle with when and how much to be paternalistic—by attempting to protect the patient against his or her own desperation and potentially poor judgment, and when to let patients choose to enter such trials, despite the potential harms. As Cynthia, an administrator, said,

> One of the most difficult issues is: is it OK to let people enroll in last resort studies? What is our responsibility for people enrolling in a study in which the prospect of benefit is very slim—almost entirely risk? But the alternative is no treatment. The study can make their last days worse! One of the hardest things is to allow people to have self-determination, and just make sure that all the information has been presented—that they're not coerced.

But if some doctors may try to pressure their patients too much, others may be too vague: Cynthia added, "Some doctors just tell such patients, 'Yeah, we've got a research study.'"

SOCIAL RISKS

A researcher may want to study rates of psychiatric problems, drug use, criminal behavior or IQ tests among a particular group—whether the Havasupai, African Americans, or gay men. But the results of such a study, if it found higher rates of problems in these populations, could be used to discriminate further against these groups. IRBs then face questions of whether to allow, change, or disapprove such protocols.

According to federal regulations, IRBs should assess potential social *benefits* of research,

but not such long-term social *risks*. The regulations stipulate that:

> 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. [31]

For example, if a researcher wanted to study rates of HIV and drug use among inner-city teenage girls, in order to help design a preventive program, the IRB may be hesitant to approve the research, fearing that there is a risk that these girls will become stigmatized and that law enforcement officials may "crack down" on this population and arrest them. The regulations dictate that the IRB should not consider such a possibility in the risk–benefit assessment. Yet the Obama Administration's proposals on IRBs posed questions about this stipulation:

> Do IRBs correctly interpret this provision as meaning that…it is not part of their mandate to evaluate policy issues such as how groups of persons or institutions, for example, might object to conducting a study because the possible results of the study might be disagreeable to them?[31]

The pediatrician and ethicist Alan Fleischman and several colleagues have felt that local IRBs *do* nonetheless at times consider broad social risks, though the regulations tell these committees not to.[32] These authors considered four realms—behavioral genetics, adolescent behavior, harm reduction, and human genetic enhancement—and concluded that IRB considerations of these social risks "sometimes create significant delays in initiating or even prevent such research."[23] These scholars oppose IRBs' considerations of these issues, since "predicting negative effects of new knowledge on populations or social policy is highly speculative and essentially political."[32] Instead, they argue, national review bodies should address these issues. Yet it is unclear what would be involved in having such national committees performing this task. After all, national committees may not be better equipped than IRBs to resolve the dilemmas that arise.[33] Committees face these issues concerning not only the four specific realms these authors discussed,[32] but more broadly.[22]

Strikingly, no prior empirical studies have examined whether IRBs do in fact consider social risks, and if so, how. I found that these categories in the regulations—"social risk," "individual risk," and "justice"—seem to be more distinct in the abstract than in the messy real world, where they can blur and become vague.

I discovered that IRBs do in fact at times oppose studies that exacerbate existing inequalities in health services. Yet this exacerbation would seem to fit the current regulatory definition of what IRBs should *not* consider. Morally, however, such a consideration may be important, suggesting that the revision of the regulation should be contemplated. Obviously, though, the specifics of when exactly IRBs should consider such risks need to be carefully assessed and decided.

Investigators may not all recognize potential social harms, which can exist even in seemingly "minimal risk" research. These interviews showed me how IRBs struggle with how to define and balance social risks, and whether and how much to do so. In practice, this category often proves related to potentially amorphous issues of stigma, vulnerability, and social inequity. After all, risks to a group can affect individuals within the group. Hence, IRBs regularly consider possible long-range social risks. Committees consider social and

psychological harms to a population (related, for instance, to stigma), and social vulnerabilities that affect groups as a whole (related to the benefits and burdens of research), but have to decide how much to do so. For example, as noted in Chapter 1, researchers had published data on high rates of schizophrenia among the Havasupai tribe. Such data about rates of mental illness in a population can increase the amount of stigma and discrimination that that group already faces. Certain cultural groups may perceive potential harms, even if the IRB might not do so. As Anthony, referring to the delicate issues surrounding human tissue samples, explained:

> Just because a sample has been de-identified from an individual standpoint doesn't mean it has been from a racial or ethnic group standpoint. There could be harm at *that* level. A group may have spiritual or worldview-related beliefs about that tissue that are much different than ours: we want bones of our ancestors returned to us, because they're not merely bones. From their perspective, it's very unpalatable that you have my blood or genes in a freezer somewhere. … So we have to expand our vision.

Consequently, while regulations explicitly state that IRBs specifically should *not* include long-term social harms as risks, some committees do so anyway, since social harms can include stigma and concerns about vulnerable populations, which the regulations do mandate IRBs to consider. Social harms can also endanger individuals who are members of the affected group.

These attitudes reflect in part the Havasupai recent lawsuit—of which many interviewees were aware, and which entailed social harm to the tribe.[34] Hence, IRBs may ignore this provision of the regulations *not* to consider "long-term" social risks, due to concerns and fears about perceptions of potential legal liability.

Risks to Vulnerable Groups

In 2004, Dan Markingson, a 26 year-old celebrity tour bus driver hoping to become an actor or screenwriter, developed schizophrenia with paranoid delusions and thoughts that he needed to murder his mother. Doctors committed him against his will to a psychiatric hospital affiliated with the University of Minnesota, and enrolled him in a study of new antipsychotic drugs. Two weeks later, they discharged him to a halfway house. His mother phoned his doctor, saying that he was getting worse, becoming suicidal, and that he did not understand the study (and should not be a participant in it). But the researchers continued. Five months later, he slit his throat and died.[35]

The FDA investigated Dan Markingson's death and did not find wrongdoing, but a bioethicist at the University, Carl Elliot, alleged otherwise: that the researchers should *not* have enrolled him in the study, and that the university was trying to protect the researchers. After Markingson's death, the Minnesota State Legislature enacted a law preventing involuntarily hospitalized psychiatric patients from participating in drug trials, unless the treating psychiatrist submits an affidavit citing the benefits to the patient; the treating psychiatrist also cannot be a researcher working on the drug trial.[36]

The reasoning behind this decision is that this population is particularly vulnerable. Patients who have been involuntarily hospitalized may not understand that they have a right to say no, or may not be mentally fit to make a reasoned decision. They also may feel coerced into

participating—perhaps believing that they will never be allowed to leave if they don't follow their doctor's recommendations. At this point in time, however, Minnesota's law, enacted because of Dan, appears unique among the 50 states.

IRBs, however, frequently wrestle with questions of who is vulnerable, and whether researchers should recruit them into studies anyway. The regulations dictate that committees "should be particularly cognizant of the special problems of research involving vulnerable populations, such as children, prisoners, pregnant women, mentally disabled persons, or economically or educationally disadvantaged persons."[37] Vulnerable individuals, lacking power, may also fail either to understand that they can readily withdraw from a study at any point, or to feel empowered to do so. They may fear angering powerful researchers. IRBs thus seek to prevent researchers from exploiting vulnerable groups, such as the poor, semi-literate black men in rural Alabama examined in the Tuskegee syphilis study. Arguably, Dan Markingson, severely psychotic because of schizophrenia, should also not have been allowed to participate in research because he both may not have sufficiently understood the study, and was involuntarily hospitalized. His rights had already been taken away. He may have felt obliged to consent in order to get treated better.

The Belmont Report and the Common Rule each suggest several differing notions of the term *vulnerability*—as related to diminished capacity, possibility of undue influence, injustice (e.g., to being unfairly burdened by a study),[38,39] and "inequitable distribution of the burdens and benefits of research participation."[40] The Declaration of Helsinki states, too, that research on a vulnerable community is only justified if the group "stands to benefit from the results."[41] In all of these documents, vulnerability is a characteristic of *populations* (i.e., it is social, not individual, in scope).[22] According to the *Oxford English Dictionary, vulnerable* means "open to temptation, persuasion…liable or exposed to disease,"[42] suggesting that both individuals *and* groups may be inappropriately recruited into a protocol and also be harmed because of it. But how IRBs in fact approach these issues has not heretofore been examined.

The regulations require that the benefits and burden of the research be distributed fairly,[36] and thus that "selection of subjects is equitable." Hence, IRBs must weigh the risks to vulnerable subjects against both the social (i.e., scientific) benefits of including these individuals, and the inequality of excluding them. Here, committees grapple with several challenges: to construct safeguards to protect vulnerable subjects, and to gauge how effective these protections need to be. Balancing possible harm to vulnerable subjects against possible scientific benefits to society can be among the most difficult decisions an IRB faces, and committees vary in how they respond to this tension.

Even studies consisting merely of interviews might harm stigmatized groups in unforeseen ways. "We try to build safeguards," Phil, a social scientist and chair, said about research on HIV, drugs, and suicide, "but know that something can go wrong."

Removed as they are from the field, IRBs find such potential dangers hard to assess—whether these harms will occur, and if so, how commonly, to what degree, and with what effects, and how to weigh all these factors. Even after IRBs erect protections, they may still worry, given lingering uncertainties. As a researcher himself, Phil thought it was important that the research proceed, but other IRBs may disagree. Additional vulnerable groups might also

warrant special protections. “Lots of other populations are special and vulnerable,” Henry, a chair, said, citing “people who are economically disadvantaged, have low levels of health, or literacy, or are in poor countries.”

The regulations do not specifically address, for instance, psychiatric patients. Dan Markingson was vulnerable, and his case highlights the challenges in researching patients with mental illness. IRB members may feel that mental health is relevant, but psychiatric disorders vary from mild to severe—from situational anxiety to severe depression and psychoses. The boundaries of a vulnerable group are also not necessarily demarcated. Not all individuals in a so-called vulnerable group may indeed be vulnerable. IRBs can differ in where they draw the line. Christopher, the physician and chair, said,

> If you’re clearly psychotic, you’re vulnerable. But if you have social anxiety disorder, I don’t think you are, and the investigator doesn’t need to have an independent psychiatrist interview each patient, and assess whether the person understands the consent and the study.

IRBs may thus struggle to decide on a case-by-case basis, defining, applying, and determining the boundaries of “vulnerability.” “It’s not cut and dry,” Christopher concluded. “You need flexibility.”

How Much Justice?

My father had insurance to pay for his care, but innumerable patients both in the United States and abroad lack such resources. Many broader questions therefore arise on IRBs concerning whether investigators and committees have responsibilities to address such ongoing health inequities when they intersect with ongoing research, and if so, how much. American bioethics has been criticized in general for overemphasizing individual autonomy at the expense of social justice and communitarian concerns.[43]

While Western European countries all guarantee certain minimum levels of insurance and care for all citizens, President Clinton’s proposals to expand health care coverage and President Obama’s Affordable Care Act have produced acrimonious battles. In the US, prevailing values have supported maximizing individual choices for those with health insurance (emphasizing individual autonomy), not guaranteeing that all patients receive coverage (which would maximize social justice). These tensions arise in research ethics as well. How much should these committees weigh not only individual, but broader *social* risks and benefits?

Martin, a researcher and chair, described, for instance, a study in which the pharmaceutical company:

> “will provide the drugs for free, but bill the insurance company for all the doctor’s visits, the time in the hospital, the CT scans, and tests. Some private insurance companies will pay. Others won’t—if it’s an experimental treatment. But then, either the drug company has to pay, or the study won’t get done. Issues then come up: what about *poor people* who don’t have insurance? Here’s a potential life-saving treatment that *only the rich* can get. The drug companies claim that they can’t otherwise afford to conduct these studies. But many of our IRB members, especially our lay members, get upset about this.”

Different opinions about these issues may therefore partly reflect members’ own backgrounds and positions on the committee. Even with the Affordable Care Act, insurers may

vary widely in whether they will cover such experimental treatments. Unfortunately, certain studies may necessitate lengthy hospital stays, for which insurance coverage is key. Hopefully, drug companies won't abuse these limitations, and seek to avoid paying for research-related expenses that they could potentially afford.

Generally, interviewees felt that they could *not* eliminate the larger health care inequities in the US and the developing world—these larger health policy issues simply lay beyond their scope. Nevertheless, IRBs may confront closely related questions—for example, whether to consider reduction or exacerbation of existing social inequities, respectively, as a social benefit or risk.

Committees thus face questions of how and to what degree to incorporate and weigh justice and injustice. When researchers want to exclude from a study patients without insurance, IRBs must decide how to respond. Including all populations equitably into a study can heighten logistical and financial costs. IRBs thus wrestle with how to balance advancing justice *vs*. facilitating research. In the US and abroad, subjects may lack health insurance, and IRBs must then decide *how much responsibility* funders and researchers have for treating medical problems that may occur during a study or afterwards as a result of the intervention. These boards encounter dilemmas in what standard to use—*how much* social justice to require, how to decide, and how much to incorporate broader justice and health inequity concerns into decisions.

Committees often try to address these ambiguities and tensions by seeking compromises. IRBs may debate, for instance, how small the criteria for exclusion can be—what to do with studies that require that participants have a high-speed Internet connection, since certain subjects may lack one. Committees may develop informal "rules," permitting exclusion of Internet-less subjects in a *pilot* study, but not in a full protocol. Charlotte explained,

> We decided our *rule of thumb* is: it's OK to exclude people for a six-month preliminary study. But for a Phase II program, a PI needs to get high-speed Internet to the participants—have them use the Internet at a clinic, or pay for it for them.

Yet even when such compromises are attempted, problems can ensue and objections can arise, based on other interpretations of the regulations. Charlotte recalled,

> A grant reviewer said, "You're still excluding a whole population of people!" The PI answered: "We're not going to be marketing this intervention to people who don't have the technology to support it. So it doesn't matter." But to exclude people flies in the face of justice!

IRBs must decide whether to compromise or make exceptions, depending on the type or extent of the study.

Committees may allow inequalities to continue partly because doing otherwise would significantly burden researchers. In predominantly white regions, for instance, IRBs face tensions concerning how ethnically diverse a sample needs to be. If the population near an institution is 98 percent white, an IRB could potentially urge or require the researcher to collaborate with researchers in other regions. But doing so imposes burdens that IRBs may not recognize. At other times, committees may allow low ethnic or racial diversity, given the obstacles to proceeding otherwise. Jack, the rural physician and chair, said:

> Occasionally, if the PI is trying to study some rare cancer, the committee says, "The tumor registry here sees one case

per year. You need to find collaborators elsewhere." But it's usually due to rarity, not diversity.

IRBs may overlook justice concerns about sampling in part because the regulations do not *clarify to what extent* IRBs should ensure or further justice. However, IRBs may consider justice only after all else—particularly the individual risks and benefits—have been assessed. IRBs may thus be unsure how much additional justice, if any, to require.

A few interviewees wondered if protections against social harms to vulnerable groups may at times go too far, impeding potentially beneficial research. Given past lapses of research ethics, sensitivities arise in studying certain vulnerable populations such as Native Americans. As Elaine explained,

> It can take months to get tribal permission, even for a really low-risk study. So a lot fewer investigators are doing research targeting Native Americans because it'll take so long to get it approved.

Wide, ongoing health disparities therefore exacerbate these dilemmas. The regulations themselves do not directly address whether researchers, funders, and IRBs have responsibilities concerning these broader social injustices, and if so, who does so and how much.

For instance, the Kennedy Krieger lead paint study sparked controversies because researchers gave subjects amounts of an intervention that were lower than the quantities known to be effective. The researchers justified their decision by arguing that the subjects would be unable to access the full treatment otherwise, and that providing and studying the medical intervention (amount of lead paint removal) at a lower dose was thus necessary. These claims remain contentious, highlighting underlying quandaries about how much effort all of us, whether as researchers, IRBs, clinicians, or taxpayers, should devote to reducing global and domestic health inequities directly *vs*. advancing science or pursing other goals.

Yet, studies may exclude patients without insurance, forcing IRBs to decide whether to disapprove such plans. Committees face dilemmas about who should pay for certain parts of these protocols, and how to present any costs to potential participants. Full disclosure may be essential, but researchers and pharmaceutical companies may not want to scare off possible subjects. As Christopher said,

> Our biggest disagreements are about who's responsible for paying for extra things. If you're in a National Cancer Institute [NCI] protocol, and get randomized to a surgical *versus* non-surgical intervention, in addition to chemo and radiation, and your insurance won't pay for the surgery because it's *experimental*, who pays for it? Should the NCI or the patient with terminal cancer pay? The researchers can tell patients, "It's going to cost $5,000. You can be in it or not." But when no one knows *how much* it will cost, there are issues. So IRBs differ. We try to come up with standard consent forms and templates, and have as much standardization as we can. But one IRB says patients should pay. Another IRB says that NCI should pay, or that it should be more explicit in the consent form: "You may be responsible for paying additional dollars. We don't know how much it is, but it could be up to $100,000." The consent forms now just say, "You will be responsible…" but not *how much* it could be. It used to say, "Please call your insurance company." But patients won't call. Some sponsors don't want patients calling up or having to pay, because subjects will drop out.

Committees confront quandaries, too, about not only direct, but *indirect* social benefits. In public health research, for example, IRBs may look for indirect benefits such as improving public health policy, since direct benefits to patients may be absent. Committees encounter questions as well of how to weigh free general medical services offered as part of a study. Laura, a lawyer and IRB administrator, described a study of tuberculosis (TB) prevalence in

the developing world,

> The study results will go to the Ministry of Health and inform policy, affecting new programs that serve people. We make sure that circuit is in place, because you're often not going to have direct benefit. By being in the study, poor populations are likely to get better medical attention for other diseases than the one being studied. A secondary benefit is that they'll be treated for free for sexually transmitted diseases [STDs] that are diagnosed, even though the study is on TB.

DEFINING MINIMAL RISK

IRBs must decide which studies are minimal risk, since the regulations for such research are different and less onerous, but making these judgments can be hard. Researchers conducting the Stanford prison, and Milgram (and as we will see, restaurant [in Chapter 11], and "fake grad student" [in Chapter 5]) studies all thought that their experiments were minimal risk. An IRB reviewed only the last of these and concurred, concluding that the researchers didn't need informed consent. Yet in each of these four cases, controversy ensued.

IRBs can vary widely in deciding if a protocol is minimal risk. In 2002, David Wendler and colleagues at the NIH surveyed 188 IRB chairs about how, in pediatric research, they applied the categories of risk in the federal regulations—minimal risk, minor increase over minimal risk, and more than a minor increase over minimal risk. These investigators asked about performing 12 procedures on an 11-year-old, and found wide variations.

Among these chairs, 27 percent thought allergy skin testing—part of routine care, and thus arguably minimal risk—was more than a minor increase in minimal risk.[44] A single car trip is part of an ordinary activity of daily life, and would thus be minimal risk; it causes death in 1/100,000 cases. Hence, any activity with that degree of risk would be minimal risk. Yet 59 percent of chairs thought that a pharmacokinetic study that had a risk of death of 1/100,000 was the *maximal* level of risk. Older chairs were more likely to see several procedures as less risky.

IRBs can encounter difficulties gauging minimal risk in part because it is relative and involves predictions about the future. Researchers' own self-assessments can be biased. As Judy, the physician and chair, said,

> The anesthesia department thinks that any anesthesia protocol is minimal risk because they have such a low rate of complications—they lose only one patient in 100,000. But none of the rest of us think that to be anesthetized is minimal risk.

Since the requirements for minimal risk research are is less burdensome, researchers may push to squeeze their studies into this category, while IRBs are cautious. Administrators may come to resent such investigator desires. "Researchers think 'expedited' means in a hurry," Cynthia, an administrator, explained. But of course, it does. Yet IRBs often see this word narrowly in procedural terms—that it refers to a category of studies—and may lose sight of the fact that the category was designed to allow for quicker reviews.

The phrase "minor increase over minimal risk" can also be ambiguous, triggering other definitional debates. Judy questioned, for example, having children spend five more minutes in a catheterization lab if they did not directly benefit:

> If three more biopsies are taken in a colonoscopy, is that a minor increase in minimal risk; or is it not allowable? We discuss that a lot. If a kid is getting a cardiac catheterization because of underlying cardiac disease, and spends five extra minutes in the cath lab getting an injection—but would already be having a cardiac catheterization for an hour and a half—is that a minor increase over minimal risk or allowable? The child is not normal. The risk is more than that of everyday life. It's really hard to know how to interpret that for sick children already undergoing that procedure. So we debate it. In the end, we decided that five extra minutes in the cath lab is OK.

Yet others might contest this conclusion. Judy feels more government input is needed concerning these definitions. "The guidelines should clarify what the criteria should be for full board review." She felt that threshold should be not more than minimal risk, "but *the risk over what risk is already happening*" [emphasis added].

Ambiguities also exist as to whether for minimal-risk studies IRBs need to review every alteration researchers make, no matter how small, after the committee has approved the project. The consensus among those I interviewed felt that IRBs do need to do so, despite consequent delays to research. Yet some chairs think this practice should be changed. In the meantime, IRBs generally say that all alterations require approval. Jeff, the social scientist and chair, said,

> If the PI is going to add one more highly standardized instrument that takes ten minutes, or realizes that a demographic is left off one of his questionnaires, does he need an amendment? In medical studies, every little change like that needs to be reviewed for safety of patients. But in a psychology study?

Certain changes in studies can pose other questions, such as definitions of when a study becomes "coercive." Jack, the rural physician and chair, said, "A sponsor just realized that it's a drag to have people sit in clinic just to hang out for blood draws, and wants to pay them $500 per day." In doubt, Jack asked the full committee.

> $500 is a lot of money. We have poor people here. That could be a little coercive. This study happens to be zero risk. But if this were a new chemotherapy study, or a Phase I study, and the risks were more severe, the money could be more coercive. I tend to feel a whole lot more comfortable if I have *all* the committee's thoughts.

These decisions may thus partly be matters of *comfort*—not wholly rational, but reflecting in part emotions, underscoring how moral decisions do not always result from entirely logical processes per se. "We have a pretty low threshold," Jack added, "for tossing things back to the full committee." As a minimal-risk study, the regulations permit him to make a decision about it quickly by himself without having to consult the entire board, which would then need to read and review the protocol, and may only meet once a month.

Given ambiguity, many "err" on the side of caution and try to be conservative. "If it's not clear," Greg, the social scientist and chair, said, "we make it expedited, rather than exempt." Such caution can, however, alienate researchers. It can also consume the IRB's own resources and time. Liza said, "IRBs just don't understand what they can expedite…to leave time for the important, challenging protocols."

Whose Daily Risk?

Since regulations define minimal risk as the risk ordinarily encountered in "daily life," questions surface of *whose* daily life. Normal, healthy volunteers, for example, face different risks than sick patients do. These questions arise in a variety of settings involving a range of potential dangers from relatively mild to more severe. Aaron, an IRB chair, cited the

recommendation of the Secretary's Advisory Committee on Human Research Protections, which says,

> "It should be the everyday life of a *normal, average person*." But I'm in my 60s—if you throw *me* on a treadmill, and run me up to twice my heart rate, there's a risk. But not for a conditioned athlete. So to me, the everyday life of *a subject* makes sense.

Moreover, even "healthy" populations range in their daily risks. Homeless inner-city street youth, for instance, even if they are healthy and drug free, may face more dangers in their lives than do middle-class suburban adolescents of the same age. As Charlotte, an IRB administrator, added:

> An intervention with street youth may not be risky to them because they're already living on the street in a dangerous environment. But if you propose it with kids living at home, it would appear risky. We argue about it, but tend to choose the norm of the population the researcher's working with. IRBs could benefit by having this articulated.

RISKS TO RESEARCHERS

IRBs also wondered whether they should extend beyond the regulations in considering risks to *researchers*, and if so, when and how much. The regulations discuss IRB responsibilities to protect subjects, not researchers.[12] Yet though not technically within its purview, a committee may try, for example, to protect student or other investigators. An additional motivation in such cases may be that the IRB is also protecting the institution—in this case, its employees—from physical or other harm. Arguably, however, such concern may be beneficial. Committees struggle with how to weigh and address these risks. Elaine said,

> One study would be exempt because the researchers are interviewing convicted sex offenders, but not recording names. The topics weren't real sensitive. It actually went to the full board because the committee was concerned about the safety to *researchers*, going to subjects' homes. The IRB was less concerned about the subjects than about the researchers, and required at least a male and a female researcher to go together.

These problems can become especially vexing with studies in politically unstable countries. Scott, an IRB director, said,

> A researcher wanted to study the trauma associated with being refugees from a war-torn country. The rebels there have been accused of genocide, conducting border raids, placing the subjects, if not the researchers, at risk. Our researchers are great, but gung-ho, Indiana Jones types. So are we going to contribute to the potential for harm befalling our own researchers? We debated, and at first rejected it. Then, after we received certain assurances that minimized the risks, we approved it.

These considerations are not called for by the regulations, but suggest how much IRBs can see themselves as overseers of research, and protectors of colleagues and institutions more broadly—how much their roles as loyal employees can blur into and affect their roles following the specific regulations. In some ways, if they had the resources, they could and perhaps should take on broader roles in certain ways—though not necessarily primarily to help the institution *per se*. Here again, tensions emerge about which wider discussion and debate are needed—of whether IRBs should go beyond the regulations, and if so, when and how much.

Committees thus wrestle with whether, when, and how much to consider social risks,

indirect social benefits, fair distribution of benefits and burdens of research, broader health inequities, and vulnerability of participants, *and* how to balance these against individual risks and benefits. IRBs also face underlying tensions concerning the degrees to which they *should* weigh these concerns, and the inherent ambiguities in interpreting the relevant terms.

These issues may be much harder to assess than individual risks and benefits, which can often be measured—a particular drug, for example, may have a 25 percent likelihood of eliminating symptoms or causing a side effect, whereas the likelihood of a study finding an effective new drug is harder to quantify. In response to the Obama Administration's questions about IRBs' interpreting regulations about social harms, IRBs appear to be taking these risks into account frequently, but wrestle with how to do so. While the Obama Administration's policymakers sought to see "social harms" and "justice" as clearly distinct, in practice these terms are closely intertwined. Stigma and widened social inequity are commonly related to social harms. At times, committees develop "rules of thumb," or compromise, or accept limitations in their ability to reduce broader health and social harms.

"Coercion," "Undue Influence," and the Question of Compensation

Dan Markingson's death revealed several problems. For instance, the researcher didn't inform the IRB of the mother's report of Dan's suicidality, which this investigator should have as an adverse event. The fact that doctors entered him into the study when he was involuntarily hospitalized also raises concern about possible coercion or undue influence.

In experiments performed by Nazis in death camps, by the US Army, and at Willowbrook,[45,46] scientists gave vulnerable subjects no choice. Federal regulations require that research minimize the "possibility of coercion or undue influence,"[12] but do not define these terms. Critics have argued that IRBs are *overly* concerned about the potential danger of undue influence, obsessing over whether small increments of money—paying a subject $50 rather than $25 for participating in a study—may unduly influence him/her. IRBs may prevent a researcher from paying, say, $50, yet lower compensation may decrease enrollment.

I once placed an ad for subjects in a magazine and accidentally left out the compensation ($30). I got no subjects. I then reran the ad with the payment and got several. Potential subjects are often busy, facing competing demands on their time. Payment can thus help incentivize participation. But money also has a range of symbolic meanings, raising sensitive concerns. I soon saw how IRBs wrestled with dilemmas of how much to pay subjects, how much is too much, and whether participants should participate only for the money or whether science should be "above" the taint of monetary concern. Some critics contend that IRBs spend too much time worrying about what relatively small amount of money to pay subjects.Yet IRB over-concern about possible undue influence can unreasonably hamper science.[47] How IRBs actually view and make decisions about these concepts, however, has not been explored. Money can distort research—for instance, at times when drug companies pay university researchers—but does money overly influence research subjects?

The Belmont Report states that coercion involves "an overt threat of harm…to obtain compliance, and offer of excessive, unwarranted, inappropriate reward"[36]—for instance, if a doctor were to tell a long-standing patient, "enter this research study or I will no longer treat you." A milder concern is potential "undue influence." More recently, OHRP has distinguished on its website that:

> "Coercion occurs when an overt or implicit threat of harm is intentionally presented by one person to another in order to obtain compliance. …Undue influence, by contrast, often occurs through an offer of an excessive or inappropriate reward or other overture in order to obtain compliance…undue influence also can be subtle. …Because of their relative nature and lack of clear-cut standards on the boundaries of inappropriate and appropriate forms of influence, investigators and IRBs must be vigilant."[48]

Coercion occurs at times in clinical psychiatry, but overall is probably relatively rare in research.[47,49,50,51] IRBs frequently end up worrying, however, about the possibility of exerting undue influence by paying subjects too much.

Recently, the philosophers Alan Wertheimer and Frank Miller have argued that IRB members mistakenly see offers of payment as coercive.[51,52,53] Yet these authors conclude, "The question as to when the offer of financial payment actually constitutes an undue influence is a topic that merits separate analysis."[51]

How much exactly should subjects get paid to be in a study? IRBs often wrangle over this question. How much should researchers pay healthy subjects to be in risky, invasive research? Should a lawyer and a McDonald's cashier get paid the same amount? Neal Dickert and Christine Grady[54] have outlined several models for determining how much to pay participants. These scholars outline three possible models and advocate a "wage payment" model—providing a low, standardized wage that could be increased for uncomfortable or other onerous tasks. They present limitations of a "market model" (based on supply and demand, and potentially offering more payment for taking on more risk), or a "reimbursement model" (covering expenses, including costs from missed work).

But, what such a wage should be, how IRBs should determine it, and how much participants actually get paid are very unclear. Online recruitment sites usually offer compensation.[55] Most journal articles, however, do not mention whether or how much participants have been compensated in the studies they describe.[56] One study gave a quantitative survey to IRB members and found they vary widely in their views.[52] Many questions remain, though, about how IRBs themselves actually make these decisions concerning studies they review, and how they perceive and experience these issues.[54,55]

As summarized in Table 3.1, IRBs wrestle to define "undue influence," and often rely on "gut feelings." Yet IRBs vary—even single IRBs shift their views from one meeting to another—reflecting underlying quandaries of whether subjects should be motivated by altruism *vs.* money.

Table 3.1. Ambiguities and Dilemmas Faced by IRBs Concerning Coercion and Undue Inducement

IRBs Struggle with Dilemmas Concerning:

• *Content*
  ◦ How much to give subjects
  ▪ Should subjects get paid differently based on their income?
  ▪ Will selection bias result?
  ▪ Is the provision of free care coercive?
  ▪ What to give subjects (e.g., cash *vs*. vouchers)?
  ▪ What types of studies
  ◦ Added challenges in several situations:
  ▪ Research on children
  ▪ Research on students
  ▪ Research in the developing world
  ◦ Whom to compensate
  ◦ When to compensate subjects
  ◦ Whether, when, and how to inform potential participants about compensation
  ◦ How to define undue influence:
  ▪ Based on "gut feelings" and "common sense"
  ▪ Can be subjective

• *Process*
  ◦ IRBs can take time to make these decisions
  ◦ Decisions often reflect compromises
  ◦ Underlying tensions arise:
  ▪ "Undue inducement" is inherently subjective and difficult to assess in others
  ▪ Questions arise of whether subjects should "volunteer" *vs*. do it for the money
  ▪ Lack of a consistent standard
  • Between IRBs
  • Even in one IRB over time

IRBs struggle with how much money is "too much," and how to decide. Determining at what

point exactly an amount becomes an "undue inducement" is hard. As Elaine said,

> Researchers were approved to pay participants $225 in a longitudinal study with follow-up interviews. They wanted to increase it to $300 due to the length of time. The chair said, "*That's* coercive." There was a compromise of $250. I was once a Research Assistant on a study, and we weren't getting people. My boss said, "We're paying them $30. Let's change it to $50."

Elaine and others misuse the term *coercion*, employing it instead of *undue influence* when no threat is involved. IRBs often use this term as a catch-all phrase for incentives that may motivate participation that might not otherwise occur.

Countless employees take jobs they don't like because of money. But should the same incentives be allowed for research subjects? "Research just seems different," Elaine felt. "It's a voluntary thing, and you don't want people to be trying just to put up with something because they're going to get paid." Yet the reasons for these differences are not always clear. The belief that research should be "voluntary" appears to reflect the notion that science should be "pure"—that everyone engaged in it should be doing it solely for the advancement of scientific knowledge. But of course this is hardly the case. IRBs may fear "undue influence" because money may thus "taint" both researchers and subjects, though for different reasons. IRBs may extend the view that researchers need to be "pure" to participants as well, raising fears of overcompensation.

Yet compensation can in fact motivate subjects; and IRBs can debate these issues at length. As Henry, the chair, said,

> We spend an inordinate amount of time on compensation levels, and whether it is adequate, or too much and coercive. We don't apply a common standard across all studies—developing countries *versus* the US; and within the US, impoverished communities *vs*. volunteers through Craigslist. …Investigators may get quite different and inconsistent advice from the committee *depending on what it feels like that day*. I don't think there's any agreement in the field. You come up with different numbers—if you think it's just to pay for people's transport *vs*. opportunity costs of being away from work.

Questions also arise of whether payment to subjects should vary based on the amount they usually earn *vs*. a single amount, regardless of their income. In the latter case, selection bias may occur, skewing the sample. As Henry continued,

> Should you compensate a radiologist much more than a laborer? Or should they get the same?…If you pay people differently, or the same, you're going to attract different groups of people, and that may cause adverse selection or targeting.

Many IRBs still determine acceptable amounts of compensation for each study on a case-by-case basis, but wide idiosyncrasies can result within and between committees.

CHALLENGES ASSESSING UNDUE INFLUENCE IN PARTICULAR SITUATIONS

Particular types of studies pose additional challenges. For instance, the provision of free care by researchers could unduly influence poor participants. But how to proceed otherwise is not always clear. As Christopher said,

> We have an 80 percent Medicaid population. …Are people participating in studies because it gives them free medicine, not because there's any real benefit to it? [They] get a free appointment and physical. Otherwise, they won't get

treated.

Yet Laura, in describing the TB study in the developing world that also provided better medical attention for other diseases, including free STD treatment, added, “It might be coercive. But do you then not do research?” She does not answer whether the research should then not be conducted. Yet I would argue that the research is vital to do, but requires very careful review.

Committees then encounter quandaries about whether and how to inform potential participants about compensation. Given the uncertainties involved, IRBs may shift their positions over time. As Dana said,

> We used to not allow [researchers] to put a monetary amount on recruitment flyers, because we thought it was a little coercive: “You’ll get $20 if you participate.” But we’ve concluded that a gift card for an hour interview is probably not particularly coercive. If you were to say, “We’ll pay $600 if we can take blood,” that might be a little coercive. So we’re trying to decide what the standard is, and how to handle that.

Committees face further challenges deciding how much to pay parents who enroll their children. Committees want to avoid the possibility of parents entering their children into studies for the money, since these youngsters may have little, if any, voice in the matter. As Jaime explained,

> Young minors should definitely not be paid in dollars. So the IRB is working on who should get the money: the parents or the child? The parent is taking the time off as the escort. There should be something for the parents as well. But you have to worry about undue inducement. Sometimes the researcher wants to pay hundreds of dollars to the parents. We’ll limit the dollar amount, and talk about what the children should get. It’s age-driven: we don’t like to put a lot of money into young adolescents’ hands. But it’s on a case-by-case basis, depending on the nature of the study…the age of the children, and what’s involved. …We polled other IRBs if they had a policy. Nobody did.

Given the lack of established policies, IRBs can vary.

These issues can become even murkier when research is conducted in the developing world. Amounts of compensation can range markedly based on established regional practices. Some projects may cover only transportation expenses—either as cash or vouchers—but this practice poses larger questions of whether it is unfair to pay in certain countries and not others. Tensions also arise between exploitation due to overly low amounts, and undue influence due to overly high rewards. Laura said, “I asked one researcher, ‘Aren’t you exploiting subjects?’ He answered, ‘It could cut both ways. If you’re offering money, it might be coercive because they are so poor.’”

Questions arise, too, when professors require students to participate in research as part of a class: these subjects receive not money or services but academic course credit. They may be given a choice as to which study to enter, yet IRBs may debate whether such participation primarily benefits the students’ education or their teachers’ careers. As Louis, an IRB chair, said,

> Students can earn that same amount of credit if they complete an extra paper. That doesn’t feel like extra credit, [but] coercion. Is it part of their practicum, or are they doing this to satisfy something for a professor?

Participation in a study may take much less time than writing a paper, which takes hours and days instead of minutes.

Conclusions

Committees struggle with defining and applying concepts of risks, benefits, undue inducement, and coercion. IRBs and researchers aim to follow principles, but are often unsure how and to what extent to do so. Compensation poses questions of whether the ideals of scientists should be extended to subjects as well. Scientists dedicate themselves to certain goals—to seeking the truth, by questioning claims about truth—disinterested in money per se, and being universal, communal and ethically neutral.[57,58] But IRBs often appear to feel that these principles should pertain not only to researchers, but to *participants* as well. In general, social groups seek to maintain "purity."[59] Therefore, IRBs may try to keep research as "pure" and unaffected by the taint of money as possible.

But adequate compensation to participants need not sully the purity of research. The larger issue of commercialization of science raises concerns, and should be confronted head on, but not prevent subjects from receiving appropriate payment. IRBs may worry excessively about compensation, misapplying the notion that scientists themselves be "pure." Academic medical centers, built on science, may hold IRBs to this standard—of being objective and unsullied. Yet while many people see science as objective, subjective elements may linger—personal beliefs and financial conflicts of interest may affect researchers' conduct, choices of studies, and interpretation of data.

Desires for objectivity may foster IRB zeal in avoiding coercion. Yet many IRBs appear to use the terms *coercion* and *undue influence* interchangeably. Though Wertheimer and Miller[51] have distinguished between these concepts—arguing that coercion involves an external threat and can often be observed by others, while undue influence is usually inherently subjective (e.g., involving the point at which someone's judgment is no longer entirely rational in following their interests)—critical questions remain.[60] In research, assessing someone else's internal state is inherently elusive. Hence, application of the term *undue* remains subjective and normative, involving questions that external observers may interpret differently—excessive according to whom, how an outsider observer is to know, who should make that determination, and on what external objective evidence? A study subject person may feel $75 is appropriate payment for filling out 1–2 hours of surveys, while an external observer may feel that that amount is excessive. Moreover, participants in a study vary in education, income, job, and personality, and consequently in how much they may be influenced—especially across diverse countries. An individual may feel that he or she is being rational, while an external observer may disagree. Hence, IRBs will no doubt continue to encounter ambiguities and variations in applying these terms.

Curiously, committees continue to misuse the word *coercion*. Linguistically, in daily discourse, IRBs may use this term to describe particular amounts of compensation (e.g., "*that's* coercive") because it is a single adjective, and hence easier to use and apply, while *undue influence* consists of two words—both an adjective and a noun, and is thus longer and more awkward and complex to apply (e.g., "*that* represents undue influence"). In trying to sway a board, members may feel that the former statement is shorter, stronger, and more effective, and carries a more sinister connotation.

These quandaries that IRBs face reflect other underlying tensions as well. Ambiguities emerge concerning the term *voluntary*, which the *Oxford English Dictionary* defines as "Of a feeling…arising or developing in the mind without external constraint; purely spontaneous in origin or character…not required or imposed, optional."[40] Yet, IRBs may feel that participation should be voluntary in the sense that subjects are not enrolling just because of the money. They may feel money can serve as *added* incentive, but that it shouldn't be the *only* incentive. When financial inducement is offered, then "voluntariness" may thus fall across a spectrum: participation may be neither "imposed" on individuals nor "purely spontaneous" but lie somewhere between these two extremes. Such a spectrum fosters IRB debates and variations.

Just as Daniel Kahneman described how, in making decisions, human beings in general irrationally overconsider risks and underemphasize benefits,[21] so, too, may IRBs. Yet these committees and others should thus be aware of this irrational but deeply human tendency, and seek to guard and fight against it.

Along these lines, as we will see, OHRP, the Institute of Medicine, or other groups could, however, produce guidelines that would help IRBs, researchers, and subjects by examining and standardizing ways of interpreting and applying these terms.