A framework for risk-benefit evaluations in biomedical research

Rid (Schulz-Baldes), A; Wendler, D

Abstract: Essentially all guidelines and regulations require that biomedical research studies have an acceptable risk-benefit profile. However, these documents offer little concrete guidance for implementing this requirement and determining when it is satisfied. As a result, those charged with risk-benefit evaluations currently assess the risk-benefit profile of biomedical research studies in unsystematic ways, raising concern that some research participants are not being protected from excessive risks and that some valuable studies involving acceptable risk are being rejected. The present paper aims to address this situation by delineating the first comprehensive framework, which is based on existing guidelines and regulations as well as the relevant literature, for risk-benefit evaluations in biomedical research.

DOI: https://doi.org/10.1353/ken.2011.0007

Annette Rid and David Wendler

A Framework for Risk-Benefit Evaluations in Biomedical Research

ABSTRACT. Essentially all guidelines and regulations require that biomedical research studies have an acceptable risk-benefit profile. However, these documents offer little concrete guidance for implementing this requirement and determining when it is satisfied. As a result, those charged with risk-benefit evaluations currently assess the risk-benefit profile of biomedical research studies in unsystematic ways, raising concern that some research participants are not being protected from excessive risks and that some valuable studies involving acceptable risk are being rejected. The present paper aims to address this situation by delineating the first comprehensive framework, which is based on existing guidelines and regulations as well as the relevant literature, for risk-benefit evaluations in biomedical research.

One of the key ethical requirements for biomedical research is that it have an acceptable risk-benefit profile (Emanuel, Wendler, and Grady 2000). The International Conference of Harmonization guidelines mandate that clinical trials should be initiated and continued only if "the anticipated benefits justify the risks" (1996). Guidelines from the Council for International Organizations of Medical Sciences state that biomedical research is acceptable only if the "potential benefits and risks are reasonably balanced" (2002). U.S. federal regulations require that the "risks to subjects" be "reasonable in relation to anticipated benefits, if any, to subjects and the importance of the knowledge" to be gained from the study (HHS 1991).

These requirements provide a general principle for risk-benefit evaluations. However, they offer little concrete guidance on how to make the judgments that they mandate. This lack of detailed guidance is hardly surprising; guidelines and regulatory documents need to address a broad range of issues in a reasonably concise way. However, the paucity of concrete guidance leads investigators, sponsors, and members of institutional

Kennedy Institute of Ethics Journal Vol. 21, No. 2, 141–179 © 2011 by The Johns Hopkins University Press

review boards (IRBs) and research ethics committees (RECs) to evaluate the risks and potential benefits of biomedical research in unsystematic ways, often based on little more than intuition alone. In one of the few empirical studies to investigate how IRB/REC members evaluate research risks and potential benefits, only 6 of 53 reviewers indicated that they use a systematic approach (van Luijn et al. 2002).

While intuition plays an important role in risk-benefit evaluations, reliance on *mere* intuition—attempting to determine the risk-benefit profile of procedures and studies based simply on how risky and beneficial they seem, without any appeal to intervening steps, analysis, or empirical data—increases the chances for mistakes. Indeed, empirical studies have found significant variation in how IRBs/RECs evaluate research risks and potential benefits (van Luijn et al. 2006; Shah et al. 2004; Lenk et al. 2004). These data raise concern that some participants are not being protected from excessive risks and that valuable studies involving acceptable risk are being rejected—or not even submitted given the uncertain outcomes of ethical review.

In recent years, scholars in bioethics have made significant progress in clarifying some of the key concepts relevant to making risk-benefit evaluations. There has been intense discussion over how to evaluate the risk-benefit profile of so-called therapeutic procedures and what role, if any, "clinical equipoise" should play in this process (Gifford 2007; London 2007; Miller and Brody 2003, 2007; Miller and Weijer 2007; Veatch 2007; Wendler and Miller 2007; Weijer and Miller 2004; Rid and Wendler 2010). The concept of "minimal risk," especially in the context of pediatric research, has sparked significant debate (Ross and Nelson 2006; Kopelman 2004; Wendler 2005; Resnik 2005; Freedman, Fuks, and Weijer 1993). Commentators have also explored whether there should be any upper limits on the research risks to which competent consenting adults may be exposed (Miller and Joffe 2009).

Although this literature has greatly improved our understanding, it focuses largely on specific aspects of risk-benefit evaluations. Lacking is a comprehensive framework that integrates these advances and offers practicable, step-by-step guidance for investigators, IRB/REC members, sponsors, and others charged with evaluating the risks and potential benefits of biomedical research studies. The present paper aims to close this gap in the literature by delineating a systematic framework for risk-benefit evaluations based on the relevant literature and on existing guidelines and regulations (National Commission for the Protection of Human Subjects

of Biomedical and Behavioral Research 1979; World Medical Association 2008; Council for International Organizations of Medical Sciences 2002; International Conference of Harmonization 1996; Council of Europe 2005; HHS 1991).[1]

THE CHALLENGE OF EVALUATING RESEARCH RISKS AND POTENTIAL BENEFITS

Consider a phase 2 study to test the efficacy of an investigational treatment for liver cancer. Laboratory data suggest that the investigational treatment may reduce tumor size more than existing treatments. However, both the investigational treatment and the procedures that are scientifically necessary for evaluating it—a series of blood draws, a CT scan, and a liver biopsy—pose risks to participants.

According to current ethical and legal guidance, this study should be approved only if investigators, sponsors, and IRBs/RECs determine that the risks are reasonable in relation to the potential benefits for participants and/or for society. Making this determination is challenging. It requires *methodological evaluation* as to whether the chosen methods are scientifically sound and so are capable of generating socially valuable results and whether the included research interventions are necessary for addressing the study question(s). It requires *empirical judgment* about how robust and relevant the available data are regarding the potential harms and benefits of the interventions. And it requires *normative evaluation* of the magnitude of the potential harms to participants, should they occur, and of how much value the collected data would have for society. Furthermore, risk-benefit evaluations go beyond a direct assessment and weighing of risks and potential benefits. Efforts to reasonably reduce the risks to research participants and to enhance the potential benefits for them and for society are typically seen as part of the evaluation process. Moreover, the risks and potential benefits of so-called therapeutic research interventions need to be evaluated in comparison to alternative treatments, if there are any. Hence, while the term "risk-benefit evaluation," strictly understood, refers to absolute judgments about risks and potential benefits, risk-benefit evaluations in biomedical research have complex comparative and normative aspects.[2] Given the number and complexity of these judgments, concern that unsystematic approaches to risk-benefit evaluations may lead to inconsistent, incomplete, and sometimes mistaken results is warranted.

The proposed framework for risk-benefit evaluations (see table) is intended to be one part of a broader ethical framework for biomedical

research (Emanuel, Wendler, and Grady 2000). The proposal describes steps for evaluating the risks and potential benefits of research studies, not algorithms or decision rules that yield one and only one verdict about the risk-benefit profile of each possible protocol. The framework thus draws on intuition and normative judgment while clarifying and guiding—based on current understanding—the role played by intuition. Our hope is that this guidance will assist investigators, IRBs/RECs, sponsors, and others who need to make these difficult judgments, helping to improve both the accuracy and the consistency of risk-benefit evaluations in biomedical research.

The proposed framework is based on several normative assumptions. The most fundamental is that biomedical research studies are ethical only if they have a reasonable risk-benefit ratio. We further assume that research is acceptable only if it has at least some social value. These assumptions seem plausible and are widely endorsed by existing guidelines and regulations. Nonetheless, commentators have questioned either or both of these assumptions (Wertheimer 2010; Sachs 2010; Rajczi 2004), and we do not attempt to respond to their arguments here.

Furthermore, the reasons why one thinks that a reasonable risk-benefit ratio is required for ethical research will have implications for how one thinks these evaluations should be conducted. For example, some commentators argue that the requirement of a reasonable risk-benefit ratio is justified on paternalistic grounds: it helps to ensure that participants' interests are protected given widespread decisional impairments in the research context (Miller and Wertheimer 2007). One implication of this view is that payment for research participation should be treated no differently in the risk-benefit evaluation than any clinical benefits participants might realize (Wertheimer 2010).

Our own view—which remains to be fully developed—is that the justification for risk-benefit evaluations goes beyond protecting the interests of research participants. The requirement for a reasonable risk-benefit ratio is justified for other reasons as well, including the need to protect the professional integrity of physician-investigators, to maintain public confidence in the research endeavor, and to ensure that societal gains in health and well-being are not won at the cost of exploiting even competent participants who agree to be exploited. Although we cannot provide a full defense of this view here, it informs the framework we propose.

A SEVEN-STEP FRAMEWORK FOR RISK-BENEFIT EVALUATIONS

Step 1: Ensure and Enhance the Study's Social Value

Risk-benefit evaluations often begin with an evaluation of the risks posed by the study. However, the proper first step is to ensure that the proposed study achieves a minimum level of social value. Studies without any social value cannot justify exposing participants to risks, no matter how low (Emanuel, Wendler, and Grady 2000). Hence, they can be rejected without further evaluation.

Despite its fundamental importance for ethical research, social value remains an underexplored concept. Socially valuable studies should collect data that are methodologically sound and address an important and unresolved scientific or clinical problem. Yet scientific value is rarely seen as a sufficient condition for social value. Reviewers also need to ensure that the information to be collected will provide at least a minimum amount of potential social benefit. In addition to evaluating the overall potential benefits of the study, reviewers should consider how those benefits will be distributed. For example, assuming the liver cancer study is scientifically sound, it is socially valuable because it addresses a serious disease affecting some of the worst-off patients (e.g., patients with chronic hepatitis and alcoholics), who could gain significantly from the new drug given the limitations of existing alternative treatments. Unfortunately, there currently are no clear criteria for evaluating whether a study passes the threshold for social value. Reviewers should therefore determine that the study has social value sufficient to justify exposing participants to at least some risks.[3] Importantly, the presence of social value does not depend on positive study results. For example, showing that the investigational liver cancer drug lacks in clinical effect is socially valuable because it prevents future patients from being exposed to unnecessary risks and may yield important insights into new options for future investigation (e.g., novel mechanisms of drug action).

Determining that a study achieves a minimum level of social value requires detailed knowledge of the research topic, as well as knowledge of the scientific methods proposed in the study. Ideally, the scientific merits of a study have been assessed before the protocol is passed on to ethical review boards. For example, to obtain funding from the U.S. government, study protocols must undergo stringent scientific review prior to being reviewed by an IRB/REC. However, prior scientific review does not always occur. In this situation, when reviewers do not have the expertise

necessary for evaluating a study's potential social benefits, they need to acknowledge their limitations and call on ad hoc committee members or expert consultants.

Once it has been determined that the proposed study reaches the threshold with respect to social value, ways should be considered to enhance its value. For example, the biopsy specimens from the liver cancer study could be made available for future research and thus increase the value of participants' contributions. Reviewers should recognize, however, that some ways of enhancing social value might increase the risks to participants, decrease the overall feasibility of the study, or require modifications of the study protocol. Storing specimens for future research, for instance, can be costly and requires adapting provisions for the consent process accordingly. Similarly, adding one or two additional passes to the liver biopsy yields more tissue for scientific testing but may increase the risks to participants. Reviewers should carefully consider how to balance these potentially competing considerations.

Step 2: Identify the Research Interventions

The second step is to identify the research interventions, as well as any procedures to be used for ensuring the safety of those interventions. Any treatments that are standard of care for the population in question need not be evaluated, unless the study context might alter their risk-benefit profile. In the liver cancer study, for example, the blood draws, CT scan, and liver biopsy are performed to test the safety and efficacy of the investigational cancer drug, and an ultrasound is necessary to safely guide the liver biopsy. All of these procedures qualify as research interventions. By contrast, the diuretics or beta-blockers that might be provided to treat participants' ascites or portal hypertension are not research interventions. Reviewers need to evaluate the risk-benefit profile of these medications only if the available data suggest that the research interventions may alter their typical risk-benefit profile. The investigational liver cancer drug, for instance, might reduce kidney function and thus decrease the clinical benefits of diuretics. This second step of identifying the research interventions is necessary for delimiting the scope of risk-benefit evaluations. It is also important because global judgments of the risk-benefit profile of a study as a whole may miss concerns raised by individual interventions within the study and are more likely to be inaccurate.

Research interventions must have the potential to yield important and nonduplicative information. Procedures that do not provide important

information pose risks for no potential social benefit and should be eliminated from the protocol. Furthermore, research interventions should help address the main scientific question(s) posed by the study. This does not preclude the possibility of allowing a small number of tangential research interventions that are relevant for addressing other issues, in particular if they are low risk. For example, investigators in the liver cancer study might propose to add a MRI scan of the head in order to investigate the early pathophysiology of liver encephalopathy. While the scan might yield important scientific information, it is not relevant to testing the safety and efficacy of the investigational cancer drug. Instead, the scan is a tangential research intervention that addresses a new study question. Allowing investigators to add a few tangential procedures can increase the efficiency of biomedical research. However, reviewers should be wary of studies that include a high number of such interventions, as well as riskier procedures that collect information on issues not directly related to the study's main scientific question(s). They should thus carefully scrutinize the risks of any tangential research interventions in the following step (step 3).

Prominent approaches to risk-benefit assessment classify the identified research interventions into two subgroups: (1) *therapeutic* interventions, such as the investigational drug in the liver cancer study, that are administered with therapeutic intent or warrant,, and (2) *nontherapeutic* interventions, such as the blood draws, CT scan, and the liver biopsy in that study, that are administered solely for the purpose of answering scientific questions (Weijer 2000; London 2007). According to these approaches, distinct ethical requirements apply to each group of research interventions. For example, "component analysis" requires that the risks of therapeutic research interventions be justified by their potential clinical benefits, whereas the risks of nontherapeutic interventions are justified by the knowledge that is expected to be gained from including those procedures in the study (Weijer and Miller 2004).

These approaches to risk-benefit evaluation are problematic. To begin with, the distinction between therapeutic and nontherapeutic interventions is not clear. An investigational drug in phase 1 testing, for example, is not administered with therapeutic intent and thus seems to be a clear example of a nontherapeutic intervention. After all, phase 1 trials evaluate drug safety, not efficacy. However, available data suggest that phase 1 drugs in oncology offer a prospect of direct clinical benefit for study participants (Miller and Joffe 2008). Thus, it would not be unreasonable for investigators to offer these drugs with therapeutic intent or warrant.

It is therefore unclear whether phase 1 oncological drugs are therapeutic or nontherapeutic procedures.

Even if the distinction between therapeutic and nontherapeutic interventions could be clarified, it is not morally relevant for evaluating the risks and potential benefits of research interventions (Wendler and Miller 2007). The goal of risk-benefit evaluations is to ensure that research interventions and studies do not expose participants to excessive risks of harm for the benefit of others. Whether a given level of research risk is excessive depends on the magnitude of the risks, the level of corresponding potential benefits for the participant, if any, and the level of potential social benefit from performing the intervention and the study. It does not depend on whether the risks result from a therapeutic or a nontherapeutic intervention.

The fundamental problem with different ethical requirements for therapeutic and nontherapeutic interventions is that they introduce different thresholds of acceptable net risk in biomedical research and thereby render risk-benefit evaluations incoherent. For example, "component analysis" allows competent participants to consent to sometimes significant risks without any compensating potential clinical benefits, as long as the risks result from a nontherapeutic intervention (e.g., a liver biopsy performed for research purposes only). Yet it does *not* allow competent participants to consent to risks, even if they are outweighed by compensating potential clinical benefits, if the risks result from a therapeutic intervention with a less favorable risk-benefit profile than available alternative treatments (e.g., a slightly less effective, first-generation liver cancer drug in a trial comparing the effectiveness of first- and second-generation treatments). Given that the main ethical concern about biomedical research is that participants are being exposed to excessive risks for the benefit of others, it is incoherent to allow significant net risks to participants for some research interventions but not for others. It follows that the risks of all research interventions in a study should be evaluated according to uniform ethical criteria.[4]

Step 3: Evaluate and Reduce the Risks to Participants

The third step is to evaluate and reasonably reduce or minimize the risks that the individual research interventions in the study, as well as any procedures necessary for ensuring their safety, pose to participants.[5] Risks are a function of two more basic components: (1) the likelihood that a harmful event or experience will occur as a result of an intervention, and (2) the extent to which the event or experience, should it occur, sets back

the individual participant's interests. For example, the liver biopsy in the liver cancer study poses a 160–733 per 100,000 risk of hemorrhage with hypotension or decrease in hemoglobin concentration that requires transfusion and/or other supportive measures, which might be considered a moderate setback to participants' interests. Risk evaluations thus require an *empirical judgment* about how likely different harms are to result from the given intervention and a *normative evaluation* of the magnitude of the respective harms, should they occur (Rid, Emanuel, and Wendler 2010). Reviewers should consider potential harms of all types—physical, psychological, social, and economic (Levine 1986)—and all degrees.

To implement this step properly, it is important to be clear regarding its goal. Research regulations focus on evaluating and reducing the risks that research interventions pose to participants. However, the ultimate goal of risk-benefit evaluations is not to protect participants from risks. The goal is to reduce the extent to which participants experience harm from participating in a research study.[6] Unfortunately, we rarely are in a position to know whether a given research procedure—for example, ingestion of the investigational liver cancer drug—will harm participants. Typically, the only way we can reduce harms to participants is to focus on and reduce the risks of harm. Recognizing that risk-benefit evaluations aim to reduce harms, not risks, is important for avoiding a common mistake.

Some commentators argue that "mere" inconvenience, burden, discomfort, or embarrassment should be excluded from risk-benefit evaluations (Levine 1986; Prentice and Gordon 2001). There are two possible grounds for this view. First, one might argue that mere inconveniences do not constitute a sufficient setback to participants' interests to be of concern. However, this argument ignores the possibility that a sufficient number of inconveniences incurred in a study could result in considerable harm, either as a result of additive or synergistic effects. Second, one might argue that "mere" inconveniences should be excluded because the requirement is to reduce *risks*, not harms that are certain to occur. For example, the inconvenience of undergoing a liver ultrasound is not a *risk* because it is a certainty for all participants. However, this argument misses the fact that the ultimate goal of the requirement to reduce risks is to reduce harms. When harms can be reduced directly, for example by excluding an unnecessary liver ultrasound from a study, this should be done as part of the present step.

Several considerations are relevant to ensuring that the risks to participants are reasonably reduced. Reviewers should eliminate research interventions that do not provide important and nonduplicative informa-

FRAMEWORK FOR RISK-BENEFIT EVALUATIONS IN BIOMEDICAL RESEARCH*

Step	*Elements*	*Example***	*Challenges*	*Open Questions**
1. Ensure and Enhance the Study's Social Value	1) Ensure the study methods are sound.	1) Yes (presumed true for present purposes).	1)This judgment requires expertise on scientific methodology.	(A) What exactly constitutes socially valuable research? Who should evaluate the potential negative social value of research studies (e.g., studies whose results can be abused)? What should be done to address it?
	2) Ensure the study passes a minimum threshold of social value.	2) The study addresses a valuable question: liver cancer is a serious disease and few treatments exist.	2) This judgment requires expertise regarding disease and existing treatments.	
	3) Enhance the knowledge to be gained from the study	3) The study could make specimens available for other research.	3) Measures to enhance social value may increase risks and/or costs.	
2. Identify the Research Interventions	1) Determine which interventions address the study question(s) or protect subjects.	1) The interventions would include investigational drug, blood draws, C, scan liver biopsy (with ultrasound).	1) Reviewers need to exclude standard treatments.	(A) What counts as a research intervention?
	2) Ensure research interventions are likely to yield important and nonduplicative information.	2) Reviewers could consider whether a liver biopsy is important for answering the study question(s).	2) Reviewers need to eliminate unimportant interventions.	

Step	*Elements*	*Example***	*Challenges*	*Open Questions**
3. Evaluate and Reduce the Risks to Participants	1) Evaluate the risks of each research intervention. 2) Reasonably reduce the risks.	1) Liver biopsy poses an 8-35 per 100,000 risk of hematothorax. 2) Reviewers could consider whether liver biopsy is safer with one-on-one monitoring.	1) Data are often limited, in particular regarding the psychological, social, and economic risks to participants. 2) Measures to reduce risks may conflict with fair subject selection, decrease social value, and/or increase costs.	A) Should risks to third parties, if any, be evaluated? By whom? What should be done to address them?
4. Evaluate and Enhance the Potential Benefits for Participants	1) Evaluate the potential clinical benefits of each research intervention. 2) Enhance the potential clinical benefits.	1) Only the investigational drug has potential clinical benefits: it may reduce cancer. 2) The study could focus on advanced liver cancer patients.	1) Data are often limited. 2) The study may conflict with fair subject selection or increase costs.	A) Should potential psychological, social, and economic benefits for participants be considered in risk-benefit evaluations? If so, how?
5. Evaluate Whether the Interventions Pose Net Risks	1) Determine whether the risks of each individual intervention exceed the intervention's potential clinical benefits (implies net risks): would an informed clinician recommend the intervention?	1) An informed clinician would recommend investigational drug (no net risks) and would not recommend blood draws, CT scan, liver biopsy (net risks).	1) This determination requires clinical expertise and data.	A) Exactly how should clinicians weigh the risks of interventions against their potential clinical benefits?

Table 1. Continued.

Step	*Elements*	*Example***	*Challenges*	*Open Questions**
6. Evaluate Whether the Net Risks Are Justified by the Potential Benefits of Other Interventions	1) Determine whether any of the interventions stand in a relation of strict scientific necessity or unity (unit of interventions).	1) Live Biopsy, blood draws, CT scan, and the investigational drug (presumed true for present purposes) do.	1) This determination requires expertise on scientific methodology	
	2) Determine whether any units of interventions pose net risks: would an informed clinician recommend the unit(s) of interventions?	2) An informed clinician would not recommend the above unit of interventions (it is stipulated that the potential clinical benefits of the investigational drug offset the drug's own risks and the risks of the blood draws and CT scan) but not all of the risks of the liver biopsy.	2) This determination requires clinical expertise and data.	A) See step 5
	1) Determine the level of cumulative net risk in the study: add any absolute, relative, indirect, and excess net risks.	1) There is a considerable level of cumulative net risks from the liver biopsy.	1) Reviewers need to cumulate different types of net risk	A) What is an appropriate method for cumulating net risks?

Step	*Elements*	*Example***	*Challenges*	*Open Questions**
7. Evaluate Whether the Remaining Net Risks Are Justified by the Study's Social Value	2) Determine whether the study's cumulative net risks fall within the general range of acceptable net risk: consider upper risk limits in research and in relevantly similar activities from the perspective of an ideal social arbiter.	2) The cumulative net risks will probably fall within the general range of acceptable net risk for research with competent participants.	2) Reviewers need to set upper limits on acceptable net risks.	B) What justifies and defines upper limits on acceptable research risk? C) How can defensible definitions of upper risk limits be implemented?
	3) Evaluate whether the given level of cumulative net risk is proportionate to the social value of conducting the study: would an ideal social arbiter recommend the study?	3) An ideal social arbiter would recommend the study (presumed true for present purposes).	3) Reviewers need to determine an acceptable relation between individual net risk and potential social benefit.	D) How does higher social value justify higher risks to participants (up to the upper limit on acceptable net risk)?

* The listed open questions are not specific to this framework; they would apply equally to any alternative approach to risk-benefit evaluations.

** The framework is applied to a hypothetical phase 2 study of an investigational treatment for liver cancer. The study involves administration of the investigational drug, a series of blood draws, a CT scan, and a liver biopsy. (It is assumed that the blood draws, scan and biopsy are scientifically necessary for testing the drug). We use the perspective of reviewers to describe the challenges associated with each step, but the same challenges apply to investigators, sponsors, and others who need to evaluate the risks and potential benefits of biomedical research studies.

tion (see step 2). Conversely, it can be appropriate to require inclusion of additional procedures to improve the safety of research interventions. In the liver cancer study, for example, the liver biopsy should be performed under ultrasound guidance to reduce the risk of bleeding and perforation of adjacent organs (e.g., gall bladder or colon). Reviewers should also probe whether relevant information can be gathered with interventions that are less risky than those proposed in the protocol. For instance, a MRI scan might be sufficient to evaluate the effects of the investigational liver cancer drug, and it would avoid the radiation exposure of a CT scan; the liver biopsy could be obtained as part of a clinically indicated procedure or investigators might be able to perform the necessary laboratory tests on existing specimens or blood samples; and so on.

Particular scrutiny should be given to the risks of any interventions that are not directly related to addressing the main study question(s). These tangential research procedures can be appropriate provided the risks are sufficiently low. For example, it would be acceptable to add a short questionnaire investigating the "therapeutic misconception" for phase 2 study participants to the liver cancer study. Do participants understand that the defining purpose of the study is to produce generalizable knowledge, regardless of whether they may personally benefit from the study? While adding such a questionnaire seems appropriate, it would not be acceptable to add a brain biopsy in order to investigate the early pathophysiology of liver encephalopathy. In addition, the overall number of tangential research interventions should not represent an excessive burden on participants. Adding a few tangential procedures can be appropriate, depending on the length of the study, but adding dozens of these interventions is not. Reviewers need to ensure that a few individual participants do not disproportionately bear the risks and burdens of biomedical research. Thus, they should eliminate any tangential interventions that pose high risks and also eliminate some if they are excessive in number. Finally, reviewers should consider modifying the study's inclusion/exclusion criteria to exclude participant groups who are at increased risk of being harmed. In the liver cancer study, for instance, excluding patients with highly vascularized cancers who have a high risk of bleeding from the liver biopsy might be justified.

Three considerations are important when evaluating and reducing the risks to participants. First, past experience with a research intervention is needed to determine what harms might result from it and to estimate how likely the harms are to occur. Yet data regarding the impact of research in-

terventions are often limited (Rid and Wendler, forthcoming). For example, a significant part of the available data on the risks of liver biopsy was collected at a time when biopsies were performed without ultrasound guidance. Reviewers must therefore carefully evaluate the strength of the available evidence, as well as its relevance for the study context under consideration. In the case of investigational drugs, reviewers should also consider whether experience with similar interventions might provide relevant evidence and judge the possible need for further preliminary research. To evaluate the risks of the investigational liver cancer drug, for example, reviewers might consult data on treatments that work through similar pathways.

While assessing the available evidence can be difficult and always requires judgment, taking existing data into account is better than relying on intuition alone. Extensive research from psychology shows that intuitive risk judgments are often biased (Tversky and Kahneman 1974; Slovic 1987; Weinstein 1989). For example, we perceive familiar risks, such as the risk of driving, to be less risky than they really are. Consideration of the available data mitigates the impact of biases and results in more informed and accurate judgments about research risks (Rid, Emanuel, and Wendler 2010).

Second, the risks of research interventions depend on who will undergo them, yet risk evaluations have to be made before the study begins. Reviewers must therefore use the study's inclusion and exclusion criteria to estimate and evaluate the risks to the average participant expected to enroll in the study. Unfortunately, determination that the risks of a research intervention are acceptable for the *average* prospective participant cannot preclude the possibility that the risks will be excessive for some individuals who are eligible to enroll in the study. Although this outcome is unavoidable, reviewers might sometimes require that investigators evaluate specific individuals prior to enrollment who predictably face excessive risks. For example, prospective participants in the liver cancer study might be asked to undergo a liver ultrasound if they suffer from a vascularized liver cancer, which increases the risk of bleeding from a liver biopsy. If an entire subgroup of prospective participants faces excessive risks, the study's inclusion/exclusion criteria should be revised to exclude this group.

Third, measures to reduce research risks can be costly, undermine the scientific validity of a study by overly limiting the data that can be collected, and compromise fair subject selection by denying some subpopulations the potential benefits of research. When delineating steps to reasonably reduce risks, reviewers should carefully balance these competing considerations.

In general, reviewers should focus on reducing serious risks. Significantly fewer trade-offs should be accepted to reduce risks that are very unlikely to occur, such as a 3 per 100,000 risk of a superficial kidney puncture from a liver biopsy or that involve harms of very low magnitude, such as the risk of getting a small bruise from a blood draw.

Step 4: Evaluate and Enhance the Potential Benefits for Participants

The fourth step is to evaluate and enhance the potential clinical benefits that the individual research interventions in the study offer participants. Importantly, this step requires judging both the likelihood and the magnitude of potential clinical benefits. Without this judgment, it is impossible to evaluate whether the level of potential clinical benefit of an intervention justifies the risks that it poses.[7]

The considerations relevant to this step are similar to those regarding the evaluation and reduction of risks. First, the data necessary for identifying the potential clinical benefits of the research interventions are often limited and thus require careful evaluation. Second, potential clinical benefits must be evaluated from the perspective of the *average* prospective participant, which implies that the potential clinical benefits may not apply to some individuals. Reviewers should consider adjusting the inclusion/exclusion criteria to focus the study on those participants for whom study participation would be most beneficial. Third, the wisdom of measures to enhance potential clinical benefits should be evaluated in light of their impact on costs, scientific validity, and fair subject selection.

While this step's exclusive focus on the potential *clinical* benefits of research interventions is consistent with current ethical and legal guidance, as well as common wisdom in bioethics (Macklin 1989; King 2000), this approach is not universally endorsed. Some commentators have argued that reviewers should factor the potential economic, social, or psychological benefits participants might realize during the study—for example, payment, praise, or feelings of altruism—in the risk-benefit calculus for a study (Sachs 2010; Wertheimer 2010; Jansen 2009). We return to this question.

Step 5: Evaluate Whether the Interventions Pose Net Risks

Some research interventions offer participants a prospect of clinical benefit such that the risks of undergoing the intervention are offset by the potential clinical benefits. For example, the risks of the investigational liver cancer drug might be outweighed by its potential to significantly reduce

tumor size. These interventions are the "win-win" cases in risk-benefit evaluations: they promote participants' clinical or health interests *and* they allow investigators to gather generalizable knowledge for the benefit of future patients. These interventions raise concern only to the extent that they promote participants' clinical interests to a lesser degree than available alternative treatments.[8]

By contrast, some research interventions offer participants little or no prospect of clinical benefit. These interventions pose "net risks"—risks of harm that are not, or not entirely, offset or outweighed by the potential clinical benefits for participants. If these risks are justified, they are justified by the social value of the information to be gained from the study. The key task of reviewers is to ensure that these net risks are not excessive. Thus, one of the most critical steps in risk-benefit evaluations is to separate the risks that participants assume for their own potential clinical benefit from the risks that they assume *solely* for the benefit of future patients. The fifth step therefore is to determine whether any of the individual research interventions in the study pose such net risks to participants.

Reviewers should consider three types of net risk, *absolute*, *relative*, and *indirect*. Absolute net risks arise when the risks of an intervention are not outweighed by the intervention's potential clinical benefits. Absolute net risks are *pure* when the research intervention poses risks without offering any potential clinical benefits for participants. This situation obtains in research interventions that offer no potential for direct clinical benefit to subjects—such as the liver biopsy in the liver cancer study, which serves to evaluate the efficacy of the investigational cancer drug but has no bearing on participants' clinical care. Research on healthy volunteers equally poses pure absolute net risks.

Absolute net risks are *impure* when a research intervention offers potential clinical benefits for participants but the benefits do not outweigh the risks. For example, research interventions that are not part of regular clinical care but nonetheless offer some prospect of clinical benefit can pose impure absolute net risks. For instance, the liver biopsy in the liver cancer study might offer some additional information for cancer staging that would generally not be obtainable in the clinical setting. However, the benefits of this information typically will not be important enough to outweigh the signficant risks of the biopsy. If the benefits did justify the risks, the biopsy would likely be part of standard medical care.

Relative net risks arise when the risks of a research intervention are outweighed by its potential clinical benefits but the intervention's risk-

benefit profile is less favorable than the risk-benefit profile of one or more available alternative treatments or diagnostic procedures. For example, investigators might propose a randomized-controlled trial to compare the effectiveness of a relatively inexpensive first-generation drug for liver cancer with an expensive second-generation drug (this assumes that the two treatments have not been tested head to head). The risks of the first-generation drug are outweighed by its potential clinical benefits, hence the drug does not pose absolute net risks. However, the first-generation drug poses relative net risks because its risk-benefit profile is in all likelihood less favorable than the risk-benefit profile of the second-generation drug.[9]

Indirect net risks arise when a research intervention itself has a favorable risk-benefit profile but the intervention diminishes the typical risk-benefit profile of other research or clinical procedures provided as part of or in parallel to the study. For example, the investigational drug in the liver cancer study might alter the risk-benefit profile of diuretics, a standard treatment for the ascites frequently associated with advanced liver cancer.

Net-risk determinations require weighing the risks of a research intervention against the intervention's potential clinical benefits and sometimes will call for comparing the risk–benefit profile of the intervention to the risk-benefit profiles of established alternative treatments or diagnostic procedures. Some commentators argue that such determinations are impossible to make because the risks and potential benefits of an intervention often affect different domains of health and thus lack a basis for comparison (Martin et al. 1995). The concepts of harm and benefit and their commensurability or comparability clearly raise deep philosophical issues. From a practical perspective, however, it is important to note that clinicians routinely make similar evaluations in the context of clinical care. To decide whether to recommend a particular treatment, clinicians consider whether the potential clinical benefits of the treatment (which might occur in one domain of health) outweigh the risks (which might be relevant to a different domain) in comparison to available alternative interventions, if any, and thus promote the patient's clinical interests. As part of this deliberation, clinicians carefully consider the available clinical data, as well as the opinion of patients and the expert professional community regarding the treatments under consideration.

Clinicians' judgments are conceptually complex and less than fully understood. Yet, at least in paradigm cases, clinicians clearly get the risk-benefit evaluations of different treatment options right. For example, clinicians who care for unconscious patients in the emergency room are

often able to evaluate whether a given treatment, such as surgery for appendicitis, is or is not in their patients' clinical interests.[10] This suggests that reviewers should adopt an informed clinician's perspective to determine whether individual research interventions pose, or do not pose, net risks. While the informed clinician's perspective will not provide reviewers with an algorithm for making net-risk determinations, it explicitly sets a normative standard for how these determinations should be made.

Under the *informed clinician test*, reviewers should ask whether a fully informed clinician who is committed solely to promoting participants' clinical interests would recommend that they undergo the intervention in question. If the clinician would recommend the intervention, it promotes participants' clinical interests and thus does not pose net risks. This implies that the intervention's risk-benefit profile is acceptable and—assuming the requirements of the previous steps are satisfied—needs no further evaluation.[11] If the clinician would be indifferent, then undergoing the intervention neither undermines nor promotes participants' clinical interests. Provided that including the intervention in the study is necessary in order to gather valuable information (step 2), this suggests that it has an acceptable risk-benefit profile. If the clinician would advise against the intervention, then it poses net risks that require further evaluation (steps 6 and 7). When applied to the liver cancer study, an informed clinician would not recommend that participants undergo the blood draws, CT scan, and liver biopsy, which—by stipulation—are performed for research purposes only. Thus, these interventions offer no potential clinical or health benefits for participants. Depending on the available evidence from preclinical testing, an informed clinician might recommend, be indifferent toward, or advise against the investigational cancer drug.[12]

Step 6: Evaluate Whether the Net Risks Are Justified by the Potential Benefits of Other Interventions

At this point, the review process has identified research interventions that pose net risks to participants and thus require further evaluation. The remaining two steps in the framework aim at assessing whether these net risks are acceptable. The net risks of some of these interventions might be justified by the potential clinical benefits of other interventions included in the same study. Some commentators reject this approach as representing a fallacy of the "package deal." They argue that allowing the potential clinical benefits of one intervention to justify the risks of a different intervention in the study might tempt investigators to add risky and unrelated

interventions to a protocol (Levine 1988; National Bioethics Advisory Commission 2001). If the risks of one intervention can be offset by the potential clinical benefits of other interventions in the same study, why not "fill up" the risks as much as the aggregate potential clinical benefits in the study allow?

This scenario raises concern that a few individual participants might disproportionately bear the risks and burdens of biomedical research (a concern we address in step 2 when we ask reviewers to evaluate whether the research interventions contribute to addressing the main scientific question(s) posed by the study). Certainly it would be problematic for investigators to conduct the liver cancer study and require that participants undergo a brain biopsy unrelated to the research in question, simply because the potential clinical benefits of the investigational cancer drug are so substantial that they "leave room" for more risks in the study. Yet while aggregating risks and potential clinical benefits across unrelated research interventions is problematic for this reason, it is a very different matter to aggregate risks and potential clinical benefits across research interventions that stand in a relation of *strict scientific necessity or unity* (Friedman, Robbins, and Wendler 2010).

In the liver cancer study, we stipulated that an estimate of the hepatocellular drug level is necessary to assess the investigational cancer drug, and it is possible that these data can be gathered only from a liver biopsy—and not, say, from a simple blood test. In this situation, the liver biopsy and the investigational drug form a single investigational unit: the drug must be either tested with the biopsy or it cannot be tested at all. Because these interventions form a single investigational unit, they can be treated as such in risk-benefit evaluations.[13] In this case, investigators are not merely adding in the biopsy to take advantage of a captive population. Rather, the study makes sense only with the biopsy included. Thus, if the drug's potential clinical benefits outweigh the risks of the drug and the risks of the biopsy together, the "package" of both interventions would pose no net risks to participants. This case of scientific necessity or unity provides one instance in which aggregating the risks and potential clinical benefits across interventions is not fallacious.

Importantly, this constrained approach to aggregating risks and potential clinical benefits across interventions precludes the possibility that risky and unrelated research procedures can be added to the protocol at will. At the same time, the approach helps to distinguish research risks that are justified by the potential clinical benefits to participants from risks that

are justified *solely* by the potential benefit to future patients. As we have pointed out, this is one of the most critical steps in evaluating the risks and potential benefits of biomedical research studies.[14]

Reviewers face two challenges when evaluating the risk-benefit profile of units of interventions. The first challenge is to determine whether a given set of interventions stands in a relation of scientific necessity or unity. Making this determination requires significant knowledge and expertise regarding scientific method. In particular, it can be difficult to judge whether there really is no alternative, less risky way to make a valid assessment of the given diagnostic or therapeutic research intervention. This judgment is relatively straightforward when the procedures under consideration pose low risks (e.g., a blood draw); it is often hard to find a research intervention that poses lower risks than those procedures under consideration. It is more difficult to judge whether a high-risk procedure, such as the liver biopsy in the liver cancer study, is *strictly* necessary to evaluate the safety and/or efficacy of some investigational intervention.

The second challenge is to evaluate whether a unit of scientifically necessary interventions poses net risks. This can require weighing a significant number of risks and potential clinical benefits against each other. In principle, however, there is no difference between evaluating whether a unit of scientifically necessary interventions poses net risks and making this evaluation about a single research intervention. Both judgments aim to determine whether undergoing an intervention, or a set of interventions, conflicts with participants' clinical or health interests. Reviewers should therefore repeat the informed clinician test in application to the investigational unit of scientifically necessary interventions and ask whether a fully informed clinician who is committed solely to promoting participants' clinical or health interests would recommend that they undergo the unit of interventions in question. If the clinician would recommend the unit of interventions, then the potential clinical benefits of the "beneficial" intervention in the unit outweigh that intervention's own risks, as well as the net risks of the intervention(s) that are scientifically necessary for testing it (even though some of the interventions that make up the unit pose net risks on their own). This implies that the unit of interventions poses no net risks to participants and thus has an acceptable risk-benefit profile. Assuming the requirements of the previous steps are satisfied, the risks and potential benefits of the interventions in the unit require no further evaluation. Of course, any net-risk interventions in the study that are not part of the unit of scientifically necessary interventions under consideration need to be further evaluated (step 7).

If the clinician would be indifferent towards the unit of interventions, undergoing it neither hinders nor promotes participants' clinical or health interests. Provided that the interventions in the unit stand in a relation of strict scientific necessity or unity and that including them in the study is necessary to gather valuable information (confirmed in step 2), this suggests that the unit of interventions has an acceptable risk-benefit profile. If the clinician would advise against the unit of interventions, the potential clinical benefits do not justify the risks. In this case, the unit poses "excess" net risks to participants that need to be further evaluated (step 7).[15] In relation to the liver cancer study, we stipulated it seems reasonable to assume that the efficacy of the investigational cancer drug cannot be evaluated without the blood draws, the CT scan, and the liver biopsy. Depending on the available preclinical data about the drug's potential clinical benefits, an informed clinician might recommend, be indifferent toward, or advise against undergoing the unit of the investigational drug, blood draws, CT scan and liver biopsy.

Step 7: Evaluate Whether the Remaining Net Risks Are Justified by the Study's Social Value

The seventh step is to evaluate whether any remaining net risks in the study—risks that are not justified by potential clinical benefits to subjects—are justified by the potential social value of the knowledge to be gained from the research. To make this evaluation, it is useful to proceed in three steps. Reviewers first should determine the level of "cumulative" net risk in the study by adding any absolute, relative, indirect, and excess net risks that have been identified in the analysis thus far. Second, reviewers should determine whether the study's cumulative net risks fall within the general range of acceptable net risk in biomedical research studies. Arguably, there are levels of net risks to individual participants that cannot be justified by even tremendous social value. If the cumulative net risks in the study clearly exceed the general limits of acceptable research risk, the study should be rejected.[16] If the cumulative net risks fall within the range of acceptable net risk, reviewers should evaluate, third, whether the given level of cumulative net risk is proportionate to the social value of the knowledge to be gained from the research.

Implementing these three steps poses several challenges. Estimating the level of net risk posed by individual interventions can be difficult. While the informed clinician test tells reviewers whether a research intervention or unit of interventions poses net risks, it does not tell them what level of

net risk is involved. It seems sensible to modify the informed clinician test such that the strength of the clinician's opposition to an intervention serves as an indicator of the level of net risk posed by an intervention or unit of interventions. Yet reviewers should remember that clinicians are experts in evaluating whether a given intervention promotes patients' clinical or health interests, not in evaluating to what extent an intervention conflicts with those interests.

Review committees also need to determine the total net risks based on the absolute, relative, and indirect net risks posed by the individual research interventions (step 5), as well as the excess net risks posed by any package(s) of interventions that are scientifically necessary to evaluating the intervention under study (step 6). For example, in the liver cancer study, reviewers must add excess net risks posed by the combination of the investigational drug, blood draws, CT scan, and liver biopsy (this assumes, as initially stipulated, that the latter three interventions are scientifically necessary for testing the investigational drug, and that the potential clinical benefits of the investigational drug does not suffice to outweigh the risks of all of these interventions); and the indirect net risks that might result from a reduced efficacy of diuretics or other procedures provided as part of or in parallel to the study. Unfortunately, there currently is no systematic approach to making these calculations.[17] Reviewers therefore need to rely on their judgment. To make these judgments as accurately as possible, reviewers should proceed under careful consideration of the available data. Moreover, they should focus on significant potential harms, such as experiencing a major hemorrhage from a liver biopsy, which are ethically more concerning than harms of lower magnitudes. At the same time, reviewers must consider the possibility that a high number of interventions involving mostly minor harms or "burdens" might pose overall considerable risks (e.g., a series of blood draws). Finally, reviewers should assess the possibility that the net risks of the different interventions have interactive effects.

The next challenge is to determine whether a study's cumulative net risks fall within the general range of acceptable net risk. Yet what defines that range remains controversial. There is widespread agreement that research risks must be strictly limited if informed consent from participants is not obtained (e.g., in research involving deception) or cannot be obtained (e.g., in research involving children or incapacitated adults). Most guidelines and regulations define the upper limit of acceptable risk in this context as "minimal" risk.[18] Yet, despite the widespread endorsement of the minimal risk

threshold, there is no widely agreed-on definition of minimal risk. The most prominent definitions hold that minimal risks should only result in a "very slight and temporary" impact on participants' health (Council of Europe 2005) or that minimal risks should be no greater than the risks of routine medical examinations (Council for International Organizations of Medical Sciences 2002) or the risks posed by activities of daily life (HHS 1991).

All these definitions are problematic in some way. A risk of serious or lasting harm can be "minimal" if the likelihood is sufficiently low. For example, a blood draw is widely—and arguably appropriately—considered a minimal risk procedure although it poses a very small risk of serious infection or permanent nerve damage. Some routine examinations (e.g., a colonoscopy in older adults) clearly pose more than minimal risks, while certain nonroutine procedures (e.g., Reiki) pose essentially no risks. And many daily life risks, such as the risks of home accidents, are unacceptably high. We should thus not use them as a standard for acceptable research risk. In addition, the risks of routine clinical examinations and daily life activities are dissimilar to net research risks because people mostly incur them for their own benefit rather than for the benefit of others. For these reasons, the minimal risk threshold, although widely endorsed, has been notoriously difficult to define and implement (Shah et al. 2004; Lenk et al. 2004).

Even less clear is whether there should be upper limits of acceptable risk in research with competent consenting participants and, if so, how that limit should be defined (Miller and Joffe 2009). With the exception of the Nuremberg Code (Annas and Grodin 1992), most existing guidelines and regulations set no explicit upper limits to risks in research with competent adults, provided the risks are reasonable in relation to the potential social benefits of the research.[19] To be sure, despite the absence of explicit regulatory guidance, the vast majority of IRBs/RECs probably would not approve research studies that involve a high risk of death or serious injury. Moreover, the fact that the public has had strongly negative responses to deaths or severe adverse events in studies that offered no potential clinical benefits for participants—for example, to the death of Ellen Roche, a healthy volunteer in a study investigating the pathophysiology of asthma (Steinbrook 2002), and to the drastic immune response of several healthy subjects in the phase 1 TeGenero trial (Suntharalingam et al. 2006)—suggests that many people would endorse absolute upper limits of acceptable research risk.[20] It also seems that no morally serious person would think it permissible to kill someone purely for research purposes, even if investigators had obtained highly scrutinized voluntary and informed consent

and the research had tremendous public health value—for example, by evaluating a very promising strategy for curing HIV/AIDS. All this suggests that there are upper limits on acceptable research risk, even in the context of highly scrutinized informed and voluntary consent.[21] Yet exactly what justifies and what defines these upper risk limits remains an open question. The historical example of Walter Reed's yellow fever experiments, which exposed healthy volunteers to a risk of developing a serious disease with an estimated mortality risk of 10 to 60 percent at the time (Lederer 2009), was probably pushing the limit. Even considering that yellow fever was a dire public health threat when the study was conducted—described as the "scourge of the American South" (Lederer 2009)—a 60 percent mortality risk for those who develop the disease seems likely beyond the absolute upper threshold of acceptable net risk in biomedical research, even when it involves fully competent adults who provide their fully informed and voluntary consent.

The main problem with delineating general risk limits in research is that doing so requires weighing net risks to individuals against potential benefits for society. What level of risk to the individual is acceptable for the benefit of others? Some commentators argue that it is impossible to make this judgment because the risks and potential benefits affect different people and thus allow for no trade-offs (Martin et al. 1995). However, while weighing individual risks against potential social benefits is difficult and raises profound philosophical questions (e.g., about the "separateness" of persons [Rawls 1971]), essentially all policy decisions depend on making these judgments. For example, to decide whether a new highway should be built near an elementary school, city planners must weigh the potential benefits to commuters against the risks to the schoolchildren and set some upper limit of acceptable risk for the project. These judgments are complex and less than fully understood. However, we do have clear normative ideals for how policy makers and public officials should approach such decisions: they should serve as "social arbiters" who (1) carefully consider the risks and potential benefits for all affected parties, ensuring that the risks to individuals are not excessive and proportionate to the benefits to them and/or society, (2) give everyone's claims fair consideration, and (3) treat like cases alike across different areas of policy. This suggests that reviewers should seek to adopt the perspective of an ideal social arbiter when considering upper limits of acceptable net risk.[22]

To provide sufficient context for these deliberations, it is sensible to evaluate upper risk limits in research in comparison to the risks posed by

nonresearch activities. For example, to determine whether the cumulative net risks in the liver cancer study fall within the range of acceptable research risk, it can be helpful to judge the 160–733 per 100,000 risk of serious hemorrhage from a liver biopsy in comparison to the risk of serious injury from playing a game of soccer for a charity fund-raiser or providing emergency assistance.[23] Importantly, for risk comparisons to be valid, reviewers must choose comparator activities that are relevantly similar to research and widely considered to be acceptable for the given population. From an ethical perspective, one of the key characteristics of research that poses net risks is that it entails assuming risks for the benefit of others. Comparator activities need to be similar in this respect. For example, charitable participation is relevantly similar and widely considered to be acceptable, even for children. To evaluate upper limits of acceptable net risk in pediatric research, reviewers might therefore use as a baseline the risks of acceptable charitable activities for children, such as a charity soccer game or car wash (Wendler 2005). The risks posed by live kidney donation (Miller and Joffe 2009) or emergency assistance (London 2006), riskier forms of acceptable "charitable participation" for adults, might be used to establish upper risk limits for research with competent adults. For example, reviewers might consider whether a liver biopsy in a healthy competent adult is more risky or less risky than donating a kidney.

Finally, the reviewers must judge whether the level of cumulative net risk in a study is proportionate to the study's social value (a determination that is always made within the constraints of the general limits of acceptable net risk). Yet it remains unclear exactly how these judgments should be made. Reviewers can and should adopt the perspective of the ideal social arbiter. Under the *idea social arbiter test*, they should ask whether a fully informed and impartial social arbiter would recommend the study in question. For example, would the social arbiter recommend the liver cancer study? If the arbiter would positively endorse the study, the value of the information to be gained from the study justifies the cumulative net risks to participants. If the arbiter would oppose the study, its social value does not justify the cumulative net risks, and the study should not be conducted. If the social arbiter would be indifferent, this equally suggests that the study should not be pursued.[24] Admittedly, the ideal social arbiter test provides less precise guidance than the informed clinician test in steps 5 and 6. In particular, the idea that the social arbiter gives everyone's claims "fair" consideration requires specification. Yet, as further work on these issues is being conducted, the ideal social arbiter

test supplies explicit, albeit general, normative standards for weighing individual net risks against potential social benefits.

A further difficulty relates to the legitimacy of making judgments about the proportionality between individual net risk and potential social benefit. As the cumulative net risks of a study increase, greater social value is required to justify them—up to the upper limit where net risks cannot be justified, however great a study's social value. While step 1 of this framework requires making no more than a threshold judgment about social value—ensuring that the study is valuable enough to merit exposing participants to any research risks—step 7 calls for judging the magnitude of a study's social value in relation to the cumulative net risks to participants. When the net risks to participants increase, reasonable people are likely to disagree about these judgments.

For example, some people might find the liver cancer study sufficiently valuable to justify its cumulative net risks because liver cancer is a serious condition that affects some of the worst-off patients and the available treatment options are limited. Other people will disagree with this verdict, arguing that the study is not important enough to justify the significant level of cumulative net risk. While reviewers have a mandate to ensure that the risks to participants are acceptable and proportionate to the study's social value, regular IRBs/RECs are not representative and transparent enough to serve as legitimate arbiters of reasonable disagreement. Reviewers should therefore recognize that the evaluation of studies involving high cumulative net risks requires a higher level of scrutiny and accountability than regular risk-benefit evaluations.[25] For example, were Walter Reed's yellow fever experiments to be conducted today, the protocol probably should undergo review by an interdisciplinary committee of national or international experts, whose opinion would be subject to extensive public input. The final opinion, as well as the rationale underlying it, probably should be made publicly available as well.[26]

BENEFITS OF THE FRAMEWORK

The proposed framework has two important benefits. The first benefit is practical. By offering systematic and comprehensive guidance for risk-benefit evaluations, the framework should enhance both the accuracy and the consistency of risk-benefit evaluations and thereby help reviewers to strike the right balance between protecting subjects from excessive risks and allowing acceptable research to proceed. Given the fundamental importance of risk-benefit evaluations for ethical research, use of the

framework has the potential to improve the evaluation of research risks and benefits compared to current practice. Rather than rely on intuition alone, investigators, sponsors, IRB/REC members, and others can follow guidance that acknowledges the crucial role of intuition and normative judgment in risk-benefit evaluations but that at the same time clarifies when and how intuition comes into play. The framework also calls attention to the key challenges associated with each step of analysis and offers possible ways to address them.

The framework's second benefit is conceptual. By offering a structure for systematic and comprehensive risk-benefit evaluations, the framework allows scholars to identify and better situate unsolved questions regarding risk-benefit evaluations in research. This helps to set the agenda for future research. For example, as we mentioned, an open question regarding step 4 of the proposed framework is whether the potential psychological, social, and economic benefits of participants should be considered in risk-benefit evaluations. It also is apparent that the fundamental first step of risk-benefit evaluations, which is to determine that a study will achieve at least a minimum level of social value, poses unsolved questions. What constitutes "socially valuable" research? Is scientific value a sufficient condition for social value? If not, what other conditions must be met for a study to offer the necessary minimum of social value? Step 7 likewise gives rise to a host of open questions. What justifies and defines upper limits of acceptable research risk? How can defensible definitions of upper risk limits be implemented? And how does "higher" social value justify higher risks to participants (up to the upper risk limit)? All these questions remain within the traditional focus of risk-benefit evaluations, which is on weighing risks to participants against the potential benefits to them and/or the potential social benefits of the research. However, the framework also makes clear that truly comprehensive risk-benefit evaluations must consider social risks from the research (Selgelid 2007, 2009; Green et al. 2006), as well as risks to third parties (Hausman 2007; Kimmelman 2005; Resnik and Sharp 2006). What should be done when it is possible that research results may be abused (e.g., in research on bioterrorism) and thus have potential "negative" social value? What should be done when the conduct of research exposes third parties to risks (e.g., in challenge studies with infectious diseases)? The proposed framework provides a clear structure for identifying and situating these open questions about risk-benefit evaluations.

CONCLUSION

Risk-benefit evaluations in biomedical research are critically important for protecting research participants from excessive research risks while allowing important research with acceptable risk to proceed. Based on current guidelines and regulations, as well as the available literature on risk-benefit assessment, the present paper delineates a comprehensive framework for risk-benefit evaluations to achieve these goals. The proposed framework provides a structure for making systematic and comprehensive risk-benefit evaluations, as well as for identifying unsolved challenges in evaluating research risks and potential benefits. Investigators, sponsors, and IRBs/RECs should use this framework to ensure that the risks to participants are not excessive in relation to the potential benefits from biomedical research studies. Scholars in bioethics should use it to structure their efforts of addressing the unsolved conceptual and normative challenges regarding one of the fundamental questions in research ethics.

The opinions expressed are the authors' own. They do not represent any position or policy of the National Institutes of Health, the U.S. Public Health Service, or the U.S. Department of Health and Human Services.

Thanks to Alex Racjzi, Frank Miller, Emily Largent, and Alan Wertheimer for helpful comments on earlier versions of this manuscript. Annette Rid gratefully acknowledges financial support from the Swiss National Science Foundation for this project.

NOTES

1. Unless noted otherwise, our analysis is based on the ethical guidance documents and research regulations we cite here.
2. One could argue that it would be more rigorous to distinguish these different aspects of risk-benefit evaluations and the different normative justifications underlying them from one another. While that approach is a prerequisite for careful conceptual and normative analysis, we adopt a practical perspective here. Our goal is to delineate practicable guidance for risk-benefit evaluations that allows investigators, sponsors, IRB/REC members, and others to systematically address all aspects related to evaluating the risks and potential benefits of biomedical research studies.
3. In some cases, it might be important to make comparative judgments about the social value of different study options. For example, the subject pool for research on very rare diseases is limited. In this situation, reviewers might want to consider what research to conduct, particularly when different study options are mutually exclusive. Is the social value of the proposed study com-

parable to other study options that address related questions? The need to make judgments about the *magnitude* of a study's social value and whether it is sufficient to justify the given level of net risk to participants, is addressed in step 7 of the framework.

4. For a detailed exposition of this critique of "component analysis" and related approaches to risk-benefit assessment, see Wendler and Miller 2007.
5. Existing guidelines and regulations typically require that the risks to participants must be "minimized." However, this requirement would prohibit most biomedical research: the safest way of minimizing the risks to participants is to stop conducting research altogether. We therefore prefer the formulation that the risks to participants should be "reasonably reduced."
6. In some cases, one might assume that risks themselves are of normative significance. For example, individuals often experience stress or anxiety as the result of knowing that they face a risk of serious harm. This is of particular concern in the context of biomedical research, where strong emphasis is placed on ensuring that subjects are aware of the risks that participation in research poses to them. However, in these cases, it is the harm of anxiety or stress that is of normative concern, not the fact of facing the risks per se.
7. This point may seem trivial, but existing regulations and guidelines often fail to specify that the likelihood of clinical benefits and their potential magnitude must be carefully evaluated. For example, the Declaration of Helsinki allows more than minimal risk research in participants who cannot give fully informed consent, provided that the research offers "a likelihood of benefit for them" (World Medical Association 2008). However, a *mere* likelihood of benefit may not be sufficient to justify more than minimal research risks in this population. For example, a very low likelihood of a small benefit, such as a 1 per 100,000 chance of not catching a cold, would probably not justify more than minimal risks. Thus, the careful evaluation of the potential clinical benefits of research interventions is a prerequisite for sound risk-benefit evaluations.
8. This statement is made solely from the perspective of risk-benefit evaluations. Interventions that promote participants' clinical or health interests and allow investigators to gather generalizable knowledge might raise other reasons for concern. For example, mutually advantageous ("win-win") transactions can be exploitative, which is a particular worry in research in developing countries (Hawkins and Emanuel 2008).
9. Relative net-risk judgments can become very complex in the context of international research, where questions related to determining the appropriate standard of care have been controversial. For example, while the Declara-

tion of Helsinki would make testing research interventions against the "best current proven intervention" the default (World Medical Association 2008), many commentators defend conceptions of the standard of care that entail less than the worldwide best treatment (e.g. London 2000; Wendler, Emanuel, and Lie 2004). Unfortunately, a discussion of how these issues relate to relative net-risk determinations is beyond the scope of the present paper.

10. Of course, physicians sometimes get these judgments wrong as well. Yet the possibility of mistakes is further evidence that risks and potential benefits can be compared at some level.
11. We wish to reemphasize our initial statement that the proposed framework for risk-benefit evaluations is one part of a broader ethical framework for biomedical research. Thus, the judgment that the risk-benefit profile of a given intervention is acceptable does not imply that the study of which it is part is acceptable. Further ethical requirements for biomedical research (Emanuel, Wendler, and Grady 2000) must be met to justify the judgment that the study as a whole is acceptable.
12. Two points of clarification about the informed clinician test. First, the test might seem to resemble the requirement of "clinical equipoise" that plays a central role in some approaches to risk-benefit assessment (Weijer and Miller 2004; London 2007). Clinical equipoise is the "honest, professional disagreement among expert clinicians" regarding which of one or more treatments, including investigational drugs, is to be preferred from the point of view of prospective study participants (Freedman 1987, 144). The informed clinician makes a similar judgment when she determines whether the individual research interventions in a study are likely to promote (or not to promote) participants' interests. Furthermore, the informed clinician makes her judgment by taking expert opinion into consideration. Yet, the key difference between the framework for risk-benefit evaluations we propose here and the alternative approaches to risk-benefit assessment we have cited (Weijer 2000; Weijer and Miller 2004; London 2007) is the normative status they attribute to the informed clinician test and the equipoise requirement, respectively. Alternative approaches regard clinical equipoise as a necessary requirement for research interventions that offer a prospect of direct clinical benefit for participants. This implies that the risk-benefit profile of these "therapeutic" research interventions must be comparable to or better than the risk-benefit profile of existing alternative treatments, if there are any. The requirement of clinical equipoise thus has normative force: it disallows "therapeutic" research interventions that pose net risks and thus sets a threshold for acceptable net risks for these procedures. By contrast, the informed clinician test is merely

an indicator of net risk in all research interventions; it does not set any risk threshold. Thus, the informed clinician test does not rule out research interventions that present net risks, including "therapeutic" procedures. The ethical acceptability of any net risks posed by research interventions is evaluated in steps 6 and 7 of the framework.

Second, the informed clinician test in its current formulation focuses on participant's clinical interests. It might therefore be seen as excluding the evaluation of research interventions that have the potential to enhance human capabilities. However, we have refrained from discussing this scenario for reasons of clarity. Research on enhancement can have important social value, if, for example, it evaluates the risks and potential benefits of enhancement interventions. Given her relevant knowledge (e.g., about physiology, pharmacology), an informed clinician should be able to judge participants' "enhancement interests." The informed clinician test could thus be modified for evaluating the risk-benefit profile of enhancement interventions.

13. A more formal presentation of the relation of strict scientific necessity or unity to research interventions would be this. Research interventions stand in a relation of strict scientific necessity or unity if 1) it would be impossible to obtain a valid assessment of research intervention A without the inclusion in the study of interventions B, C, . . . , X, and 2) there is no alternative, less risky way to evaluate A other than by including B, C, . . . , X. Importantly, the relation of scientific necessity is transitive across interventions. If intervention B is necessary to evaluate intervention A, and B will provide the necessary information only if participants undergo intervention C, C would also qualify as scientifically necessary to evaluate the experimental intervention. For example, a CT scan of the liver might be necessary to test the investigational cancer drug, but the scan can only yield valid data with a contrast agent. In this situation, both the CT scan and the contrast agent stand in a relation of scientific necessity or unity with the investigational drug for liver cancer.
14. One might expect that the constrained approach to aggregation is primarily relevant for research with participants who cannot consent (Friedman, Robbins, and Wendler 2010), where net risks are strictly limited. Indeed, this is the context in which constrained aggregation probably has the most significant practical implications. For example, according to current regulations and guidelines, the liver cancer study could not enroll patients who suffer from liver cancer as well as alcohol dementia because the liver biopsy clearly exceeds the level of acceptable "minimal" risk in this population. By contrast, the study could enroll these patients on the constrained aggregation view if the biopsy is scientifically necessary for testing the investigational liver

cancer drug and the potential clinical benefits of the drug outweigh the risks of both the drug and the biopsy. Nonetheless, constrained aggregation can have important implications for research with competent participants as well. By justifying part or all of the risks of net-risk interventions with the potential clinical benefits of another intervention in a study, constrained aggregation influences the level of net risk that needs to be justified by the social value of the study. For example, if the risks of the liver biopsy can be partially or fully justified by the potential clinical benefits of the investigational cancer drug, a lower level of "cumulative" net risks (see step 7) needs to be weighed against the potential social benefits of conducting the study.

15. An alternative way of making these judgments would be to simply ask whether the aggregate potential clinical benefits in the package of scientifically necessary interventions justify the aggregate risks of the package. However, we prefer the current formulation since it explicitly recognizes that one or more interventions in the investigational unit pose net risks to participants.
16. This statement assumes that all measures to reasonably reduce the risks to participants have been taken (step 3) and that, by implication, the study cannot be further modified to reduce those risks.
17. To be precise, there currently is no method for adding net risks of one type or for adding net risks of different types. It would probably be feasible to cumulate pure absolute net risks by using a modified version of the SERR *(*systematic evaluations of research risks*)* method for implementing upper thresholds of acceptable research risk (Rid, Emanuel, and Wendler 2010). Roughly, this modified SERR approach might consist of the following steps: 1) decluster the pure (absolute) net risks into a complete list of potential harms and associated likelihoods; 2) classify the potential harms by magnitude using the SERR harm scale; and 3) add the likelihood estimates of all potential harms of the same magnitude. Yet this modification of SERR could only be used to cumulate pure absolute net risks. More work on how to determine the total level of net risk in a study is needed.
18. U.S. federal regulations also allow a "minor increase over minimal" risk in certain pediatric research studies (HHS 1991).
19. The Nuremberg Code, delivered as part of the Nuremberg military tribunals' verdict on the Nazi concentration camp experiments, stipulates that "no experiment should be conducted where there is an *a priori* reason to believe that death or disabling injury will occur; except, perhaps, in those experiments where the experimental physicians also serve as subjects" (Annas and Grodin 1992). This statement appropriately directs attention to the sometimes serious risks associated with research. However, it does not specify a likelihood

threshold for when a risk of serious injury is acceptable. After all, the mere presence of a risk of serious harm, such as a risk of death, does not make a research intervention unacceptable. It is only when the likelihood of death is sufficiently high that mortality risks become unacceptable. For example, a blood draw is one of the paradigm examples of a "minimal" risk research procedure, although it poses an exceedingly remote risk of death from infection.

20. A note of caution is warranted here. It seems likely that the negative public response to these cases rested on a common mistake in risk-benefit evaluations, namely that of evaluating the risks and potential benefits of a study in hindsight. However, as Henry Beecher famously noted, "An experiment is ethical or not at its inception. It does not become ethical post hoc" (Beecher 1966). Thus, the fact that a study seriously harmed some participants does not necessarily imply that the study exposed participants to excessive risk. Indeed, it is likely that some serious harms will eventually occur in the context of acceptable research (that is, research with a very low risk of serious harm). Nonetheless, the strongly negative public response to serious injury or death—in particular in "nonbeneficial" research studies—lends support to the common intuition that net risks in research should be limited, even if this means that therefore some highly valuable research cannot be conducted.
21. We recognize that more argument is needed to properly support this claim. The only paper dedicated to this question—to our knowledge—argues that the uncertainty of potential social benefit from any particular research study calls for prudence in exposing participants to substantial net risks (Miller and Joffe 2009). Yet this line of argument does not exclude exposing participants to very high risks if the potential social benefits of a given study are substantial and almost certain to be materialized. As we have mentioned, our own view is that risk-benefit evaluations must be grounded in, among other things, the need to ensure that societal gains in health and well-being are not secured at the cost of exploiting even competent participants who agree to be exploited. This view—which still awaits a full defense—would support limits to research risks even if a study is almost certain to yield results of tremendous public health benefit.
22. Thanks to Bob Goodin for suggesting this term.
23. Risk comparisons do not only provide a context for evaluating upper limits of acceptable net risk in research. They also promote consistent risk judgments across activities in different realms of life (consistent with the ideal social arbiter perspective) and allow reviewers to appeal to risk judgments made outside the research context (Rid, Emanuel, and Wendler 2010).

24. The ideal social arbiter test requires a positive endorsement of the study; indifference vis-à-vis the study is not sufficient for establishing an acceptable level of cumulative net risk. This test criterion reflects the conviction that reviewers should be certain that the study's cumulative net risks are justified by the social value of the information to be gained rather than merely neutral on this question.
25. The following ideas about how to proceed in the case of reasonable disagreement about risk-benefit evaluations are inspired by Norman Daniels and James Sabin's "accountability for reasonableness" approach to addressing reasonable disagreement regarding the allocation of scarce resources for medical care (Daniels and Sabin, 1997).
26. It is now apparent that the present framework for risk-benefit evaluations shares key features with the net-risks-test approach coauthored by one of us (David Wendler; see Wendler and Miller 2007). However, the present framework differs from the net-risks test in at least three important ways. First, unlike the net-risks test, it provides a *comprehensive* framework for risk-benefit evaluations that offers practicable guidance for investigators, sponsors, IRB/REC members, and others. Second, the present framework allows for the constrained aggregation of risks and potential clinical benefits (step 5), which the net-risks test explicitly excludes (Wendler and Miller 2007). Third, unlike the net-risks test, the present framework does not evaluate the net risks of each individual intervention in relation to the social value of including that particular intervention in the study. Instead, it proceeds directly to weighing the cumulative net risks in the study against the social value of the information to be gained from the research. We now think it is difficult to judge the social value of including individual research interventions without considering the broader study context, and thus we proceed directly to evaluating a study's cumulative net risks in light of its social value.

REFERENCES

Annas, George J., and Grodin, Michael A., eds. 1992. *The Nazi Doctors and the Nuremberg Code*: New York: Oxford University Press.

Beecher, Henry K. 1966. Ethics and Clinical Research. *New England Journal of Medicine* 274 (24): 1354–60.

Council for International Organizations of Medical Sciences. 2002. *International Ethical Guidelines for Biomedical Research Involving Human Subjects* Geneva: Council for International Organizations of Medical Sciences.

Council of Europe. 2005. *Additional Protocol to the Convention on Human Rights and Biomedicine, Concerning Biomedical Research*. Strasbourg, Council of Europe.

Emanuel, Ezekiel J.; Wendler, David; and Grady, Christine. 2000. What Makes Clinical Research Ethical? *JAMA* 283 (20): 2701–11.

Daniels, Norman, and Sabin, James E. 1997. Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers. *Philosophy and Public Affairs* 26 (4): 303–50.

Freedman, Benjamin. 1987. Equipoise and the Ethics of Clinical Research. *New England Journal of Medicine* 317 (3): 141–45.

Freedman, Benjamin; Fuks, Abraham; and Weijer, Charles. 1993. In Loco Parentis: Minimal Risk as an Ethical Threshold for Research upon Children. *Hastings Center Report* 23 (2): 13–19.

Friedman, Alexander; Robbins, Emily; and Wendler, David. 2010. Which Benefits of Research Participation Count as "Direct"? *Bioethics*. DOI: 10.1111/j.1467–8519.2010.01825.x.

Gifford, Fred. 2007. So-Called "Clinical Equipoise" and the Argument from Design. *Journal of Medicine and Philosphy* 32 (2): 135–50.

Green, Shane K.; Taub, Sara; Morin, Karine; and Higginson, Daniel. 2006. Guidelines to Prevent Malevolent Use of Biomedical Research. *Cambridge Quarterly of Healthcare Ethics* 15 (4): 432–39; discussion 439–47.

Hausman, Daniel M. 2007. Third-Party Risks in Research: Should IRBs Address Them? *IRB* 29 (3): 1–5.

Hawkins, Jennifer S., and Emanuel, Ezekiel J., eds. 2008. *Exploitation and Developing Countries: The Ethics of Clinical Research*. Princeton, NJ: Princeton University Press.

HHS (U.S. Department of Health and Human Services). 1991. Protection of Human Subjects. 45 CFR 46.

Holm, Soren, and Harris, John. 2009. The Standard of Care in Multinational Research. In *The Oxford Textbook of Clinical Research Ethics*, ed. Ezekiel J. Emanuel, Christine Grady, Robert A. Crouch, Reidar Lie, Franklin Miller, and David Wender, pp. 729–37. New York: Oxford University Press.

International Conference of Harmonization. 1996. *ICH Harmonized Tripartite Guideline: Guidelines for Good Clinical Practice*. Geneva: International Conference of Harmonization.

Jansen, Lynn A. 2009. The Ethics of Altruism in Clinical Research. *Hastings Center Report* 39 (4): 26–36.

Kimmelman, Jonathan. 2005. Medical Research, Risk, and Bystanders. *IRB* 27 (4): 1–6.

King, Nancy M. 2000. Defining and Describing Benefit Appropriately in Clinical Trials. *Journal of Law, Medicine, and Ethics* 28 (4): 332–43.

Kopelman, Loretta M. 2004. Minimal Risk as an International Ethical Standard in Research. *Journal of Medicine and Philosophy* 29 (3): 351–78.

Lederer, Susan E. 2009. Walter Reed and the Yellow Fever Experiments. In *The Oxford Textbook of Clinical Research Ethics*, ed. Ezekiel J. Emanuel, Christine Grady, Robert A. Crouch, Reidar Lie, Franklin G. Miller, and David Wendler, pp. 9–17. New York: Oxford University Press.

Lenk, Christian; Radenbach, Katrin; Dahl, Mathias; and Wiesemann, Claudia. 2004. Non-Therapeutic Research with Minors: How do Chairpersons of German Research Ethics Committees Decide? *Journal of Medical and Ethics* 30 (1): 85–87.

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

———. 1988. Uncertainty in Clinical Research. *Journal of Law, Medicine, and Ethics* 16 (3–4): 174–82.

London, Alex J. 2000. The Ambiguity and the Exigency: Clarifying "Standard of Care" Arguments in International Research. *Journal of Medince and Philosphy* 25 (4): 379–97.

———. 2006. Reasonable Risks in Clinical Research: A Critique and a Proposal for the Integrative Approach. *Statistics in Medicine* 25 (17): 2869–85.

———. 2007. Two Dogmas of Research Ethics and the Integrative Approach to Human-Subjects Research. *Journal of Medicine and Philosophy* 32 (2): 99–116.

Macklin, Ruth. 1989. The Paradoxical Case of Payment as Benefit to Research Subjects. *IRB* 11 (6): 1–3.

Martin, Douglas K.; Meslin, Eric M; Kohut, Nitsa; and Singer, Peter A. 1995. The Incommensurability of Research Risks and Benefits: Practical Help for Research Ethics Committees. *IRB* 17 (2): 8–10.

Miller, Franklin G., and Brody, Howard. 2003. A Critique of Clinical Equipoise: Therapeutic Misconception in the Ethics of Clinical Trials. *Hastings Center Report* 33 (3):19–28.

———. 2007. Clinical Equipoise and the Incoherence of Research Ethics. *Journal of Medicine and Philosophy* 32 (2): 151–65.

Miller, Franklin G., and Joffe, Steven. 2008. Benefit in Phase 1 Oncology Trials: Therapeutic Misconception or Reasonable Treatment Option? *Clinical Trials* 5 (6): 617–23.

———. 2009. Limits to Research Risks. *Journal of Medical Ethics* 35 (7): 445–49.

Miller, Franklin G., and Wertheimer, Alan. 2007. Facing Up to Paternalism in Research Ethics. *Hastings Center Report* 37 (3): 24–34.

Miller, Paul B., and Weijer, Charles. 2007. Equipoise and the Duty of Care in Clinical Research: A Philosophical Response to Our Critics. *Journal of Medicine and Philosophy* 32 (2): 117–33.

National Bioethics Advisory Commission. 2001. *Ethical and Policy Issues in Research Involving Human Participants: Report and Recommendations of the National Bioethics Advisory Commission*. Bethesda, MD: National Bioethics Advisory Commission.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. 1979. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Washington, DC: Government Printing Office.

Rajczi, Alex. 2004. Making Risk-Benefit Assessments of Medical Research Protocols. *Journal of Law, Medicine, and Ethics* 32 (2): 338–48.

Rawls, John. 1971. *Theory of Justice*. Cambridge, MA: Harvard University Press.

Resnik, David B. 2005. Eliminating the Daily Life Risks Standard from the Definition of Minimal Risk. *Journal of Medical Ethics* 31 (1): 35–38.

Resnik, David B., and Sharp, Richard R. 2006. Protecting Third Parties in Human Subjects Research. *IRB* 28 (4): 1–7.

Rid, Annette; Emanuel, Ezekiel J.; and Wendler, David. 2010. Evaluating the Risks of Clinical Research. *JAMA* 304 (13): 1472–79.

Rid, Annette, and Wendler, David. A Proposal and Prototype for a *Research Risk Repository* to Improve the Protection of Research Participants. *Clinical Trials* (revise and resubmit).

———. 2010. Risk-Benefit Assessment in Medical Research: Critical Review and Open Questions. *Law, Probability and Risk* 9 (3–4): 151–77.

Ross, Lanie F., and Nelson, Robert M. 2006. Pediatric Research and the Federal Minimal Risk Standard. *JAMA* 295 (7): 759; author reply 759–60.

Sachs, Benjamin. 2010. The Exceptional Ethics of the Investigator-Subject Relationship. *Journal of Medicine and Philosophy* 35 (1): 64–80.

Selgelid, Michael J. 2007. A Tale of Two Studies: Ethics, Bioterrorism, and the Censorship of Science. *Hastings Center Report* 37 (3): 35–43.

———. 2009. Governance of Dual-Use Research: An Ethical Dilemma. *Bulletin of the World Health Organization* 87 (9): 720–23.

Shah, Seema; Whittle, Amy; Wilfond, Benjamin; Gensler, Gary; and Wendler, David. 2004. How Do Institutional Review Boards Apply the Federal Risk and Benefit Standards for Pediatric Research? *JAMA* 291 (4): 476–82.

Slovic, Paul. 1987. Perception of Risk. *Science* 236 (4799): 280–85.

Steinbrook, Robert. 2002. Protecting Research Subjects: The Crisis at Johns Hopkins. *New England Journal of Medicine* 346 (9): 716–20.

Suntharalingam, Ganesh; Perry, Meghan R.; Ward, Stephen; Brett, Stephen J.; Castello-Cortes, Andrew; Brunner, Michael D.; and Panoskaltsis, Nicki. 2006. Cytokine Storm in a Phase 1 Trial of the Anti-CD28 Monoclonal Antibody TGN1412. *New England Journal Medicine* 355 (10): 1018–28.

Tversky, Amos, and Kahneman, Daniel. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185 (4157): 1124–31.

van Luijn, Heleen E., Aaronson, Neil K.; Keus, Ronald B.; and Musschenga, Albert W. 2006. The Evaluation of the Risks and Benefits of Phase II Cancer Clinical Trials by Institutional Review Board (IRB) Members: A Case Study. *Journal of Medical Ethics* 32 (3): 170–76.

van Luijn, Heleen E.; Musschenga, Albert W.; Keus, Ronald B.; Robinson, Walter M.; and Aaronson, Neil K. 2002. Assessment of the Risk/Benefit Ratio of Phase II Cancer Clinical Trials by Institutional Review Board (IRB) Members. *Annals of Oncology* 13 (8): 1307–13.

Veatch, Robert M. 2007. The Irrelevance of Equipoise. *Journal of Medicine and Philosophy* 32 (2): 167–83.

Weijer, Charles. 2000. The Ethical Analysis of Risk. *Journal of Law, Medicine, and Ethics* 28 (4): 344–61.

Weijer, Charles, and Miller, Paul B. 2004. When Are Research Risks Reasonable in Relation to Anticipated Benefits? *Nature Medicine* 10 (6): 570–73.

Weinstein, Neil. 1989. Optimistic Biases about Personal Risks. *Science* 246:1232–33.

Wendler, David. 2005. Protecting Subjects Who Cannot Give Consent: Toward a Better Standard for "Minimal" Risks. *Hastings Center Report* 35 (5): 37–43.

Wendler, David, and Miller, Franklin G. 2007. Assessing Research Risks Systematically: The Net Risks Test. *Journal of Medical Ethics* 33 (8): 481–86.

Wendler, David; Emanuel, Ezekiel J.; and Lie, Reidar. 2004. The Standard of Care Debate: Can Research in Developing Countries Be Both Ethical and Responsive to Those Countries' Health Needs? *American Journal of Public Health* 94 (6): 923–28.

Wertheimer, Alan. 2010. *Rethinking the Ethics of Clinical Research: Widening the Lens*. New York: Oxford University Press.

World Medical Association. 2008. Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects (first version adopted 1964). http://www.wma.net/en/30publications/10policies/b3/index.html.