6

The Integrative Approach to Assessing and Managing Risk

6.1 Reconciling Social Value and Equal Concern

The previous chapter presented an indirect defense of the egalitarian research imperative. It was indirect in the sense that it highlighted and then criticized a network of views, some tacit and some explicit, against the background of which it appears almost true by definition that research with human participants poses a fundamental and inescapable moral dilemma. Undermining this network of assumptions clears conceptual room for an approach to risk assessment and management that creates the conditions under which free and equal people, pursuing diverse life plans in a free society, can see research participation as a viable avenue for advancing the common good.

The goal of the present chapter is to demonstrate, in concrete and operationally meaningful terms, how it is possible to reconcile the imperative to support research that advances the common good with the imperative to respect the status of the stakeholders in that undertaking as free and equal persons. What I refer to as the *integrative approach* articulates the conditions necessary to reconcile or integrate these goals.

The integrative approach is grounded in the same concern for the basic or generic interests of individuals that defines the generic interests conception of the common good (§4.5) and that motivates the egalitarian research imperative (§4.7). It is in virtue of this common focus and shared normative foundation that the requirements of the integrative approach are not extrinsic side constraints on research. Instead, they are integral and enabling components of a system of social cooperation in which free and equal persons can advance the common good with credible public assurance that their status as free and equal will not be compromised in the process.[1]

[1] The role of prospective review of research before bodies of diverse representation in providing this credible public assurance is the subject of chapter 7.

For the Common Good. Alex John London, Oxford University Press. © Oxford University Press 2022.
DOI: 10.1093/oso/9780197534830.003.0006

In §6.2 I show that pursuit of the egalitarian research imperative that I outlined in chapter 4 entails what I call the principle of equal concern. In §6.3 I present a high-level overview of the three operational criteria that the integrative approach uses to give concrete content to the principle of equal concern. In §6.4 I show how the integrative approach uses insights similar to those in Benjamin Freedman's clinical equipoise to articulate operationally meaningful practical tests for one of these operational criteria. This demonstrates how the integrative approach gives content to what I called the template for the appeal to uncertainty from §5.4 and how it reconciles the social value requirement with concern for participant welfare. In §6.5 I extend this argument to what is arguably the most difficult case to justify for approaches of this kind, namely, research designs that use response-adaptive randomization (a design in which the probability that patients are allocated to various interventions is dynamically adjusted throughout the course of the trial in light of outcomes that are observed in the study's various arms). In §6.6 I show how the integrative approach allows socially valuable research to continue to the point where it is most likely to alter the practice of stakeholders without violating a series of compelling ethical requirements.

The argument presented in these first five sections demonstrates that the integrative approach can reconcile the social value requirement with a set of intuitive and important ethical values. In §6.6 I clarify some of the criteria that should be used to evaluate competing frameworks for assessing and managing risk in the research context, and in §6.7 I highlight some ways in which the integrative approach differs from competing frameworks that also appeal to uncertainty. I conclude with some brief remarks about the distinct advantages of this approach over alternatives that reject the appeal to uncertainty and instead frame the problem of risk in research as an exercise in risk-benefit analysis.

6.2 Two Requirements of the Egalitarian Research Imperative

6.2.1 Social Value and the Public Purpose of Research

We saw in the last chapter that frameworks for risk assessment and management in orthodox research ethics are typically grounded in the moral obligations of medical professionals. This grounding contributes to the idea

that the fundamental problem related to risk in research ethics is to reconcile the individual clinician's duty of personal care with the researcher's utilitarian duty to improve the lives of future persons. This framing is understandable for historical reasons, given that research scandals that helped to shape the field often involved clinicians using prerogatives they enjoyed in light of their role as caregiver to conduct research that was antithetical to that role (§2.4). However, this framing creates the appearance of an irreconcilable conflict between the individual decision-maker's pursuit of the epistemic goals of research and a very plausible moral principle that we called the principle of concern for welfare (§5.4):

Concern for Welfare: It is impermissible to knowingly expose a person to interventions, practices, or procedures, that are known or credibly believed to be worse than another available option.

The integrative approach rejects both dogmas of research ethics discussed in the last chapter: it is not grounded in the narrow moral obligations of particular professional roles and it does not presuppose that research is an inherently utilitarian activity. It also rejects the presumption that it is sufficient to think of research in functional terms, as a set of goals and ends that structure the discrete interactions of a set of basically private parties.

Instead, the integrative approach recognizes that research is a scheme of social cooperation that serves a public purpose grounded in considerations of justice. One such consideration of justice concerns the claims that community members have on the goals and ends that are advanced by the research enterprise.

Following the egalitarian research imperative, the public purpose of research is to generate the knowledge necessary to bridge gaps in the capacity of the basic social institutions of a community—such as its system of public health and clinical medicine—to safeguard and advance the basic interests of that community's members. Research programs that satisfy this condition have a strong, prima facie claim to social value. We captured this idea in a formulation of the social value requirement that reflects the content of the egalitarian research imperative:

Social Value Requirement: Research with human participants is only justified if it is reasonably expected to generate the knowledge necessary to develop interventions, policies, practices, or other advances that will enable that

community's basic social structures (such as its health-related institutions) to more effectively, efficiently, or equitably safeguard and advance the basic interests of its constituent members.

When research takes place in a community, and it is designed to generate information that is needed to expand the capacity of that community's basic social institutions to safeguard and advance the basic interests of that community's constituent members, then it has a strong claim to represent a just division of social labor and a just use of scarce human and material resources.

6.2.2 The Principle of Equal Concern

Unlike orthodox research ethics, the integrative approach recognizes that research is a scheme of social cooperation involving the collaboration of many different parties, often extended over long periods of time. In order to be consistent with principles of justice, this must be a voluntary scheme of social cooperation in which participants have credible social assurance that in taking on the purpose of advancing the common good they will not be subject to arbitrary treatment, including antipathy or abuse, exploitation, domination, or other forms of unfair treatment.

Following the egalitarian research imperative, in producing socially valuable information this scheme of social cooperation must respect the status of stakeholders—including study participants—as free and equal. This commitment is captured in the following principle:

Principle of Equal Concern: As a necessary condition for ethical permissibility, research with humans must be designed and carried out so as not to undermine the standing of any research participant as the moral and political equal of their compatriots, by either knowingly compromising participant basic interests or by showing less care and concern for their basic interests than the interests of those the research is intended to serve.

The integrative approach to risk assessment and management seeks to reconcile these two requirements by providing a framework that ensures that studies respect the status of participants as free and equal and have a strong, prima facie claim to generating social value. This is a prima facie claim because factors that affect the social value of a study range beyond the set of

conditions that determine whether the study is designed on terms that respect the status of participants as free and equal. In order for this prima facie claim to be fully substantiated, additional information may be required about the relationship between the study; the needs of the target community; the ability of that community to use study results to expand the capacity of its health systems without additional assistance; whether necessary assistance is available, if required; and whether there are other ways of addressing those needs that are more effective, efficient, or equitable. Such additional questions are addressed in chapter 9.

The integrative approach thus articulates the conditions necessary for discharging the egalitarian research imperative. In particular, that imperative requires that communities foster research, conceived of as a voluntary scheme of social cooperation that advances the common good. Undertakings that are designed and conducted on terms that satisfy the principle of equal concern create a foundation for voluntary participation in this scheme of social cooperation by providing credible public assurance that in volunteering to advance the common good participants are not sanctioning practices that would make them subject to arbitrary treatment, subjugation, domination, or abuse.

6.2.3 Justice and the Common Good

We can also frame the problem the integrative approach is designed to solve in terms of the conditions we articulated in §4.2.2 on appeals to the common good. Because sickness, injury, and disease threaten the physical, intellectual, affective, or social capacities needed to formulate, pursue, and revise a life plan, they threaten to compromise abilities that are fundamental to the status of individuals as the moral and political equals of their compatriots. This threat to the basic interests of individuals is sufficient to satisfy the triggering condition (TC) for the normative claim (NC). This is a circumstance in which it is permissible to ask individuals to risk, sacrifice, alter, or limit ends or goals that are part of their individual life plan—their personal interests—in the service of an effort to secure the basic interests of others. On the basic or generic interests conception of the common good this means that there are strong moral and political reasons to use social resources and social authority to create cooperative arrangements that promote opportunities for individuals to take up, as part of their personal life plan or as a personal project, activities that advance the basic interests of others.

However, because every community member has an equal claim on the basic institutions of their society to use the best practices available to safeguard and to advance their basic interests, efforts to advance the common good must be consistent with the practical constraint (PC) that the means used must not themselves conflict with or subvert the common good. The principle of equal concern represents the condition necessary to ensure that social efforts to advance the common good do not violate this practical constraint.

The approach to risk that I'm outlining here is "integrative" in the sense that it strives to address social problems by using social resources and social authority to create opportunities for some community members to exercise their shared basic capacities for agency and welfare by taking on personal projects that have as their goal securing these same basic capacities for others (§4.5). At the end of the last chapter we saw that careers in medicine, including careers in medical research, are easily conceived of on these terms (§5.11). They provide a substantive outlet through which individuals can develop and pursue a range of personal talents and abilities while also advancing the common good.

Adhering to the principle of equal concern allows society to treat research participation as an avenue through which its members might advance the common good without compromising their standing or status. To say that being a researcher and being a study participant should each be seen as an outlet for advancing the common good is not to say that the latter offers participants the same kind of outlet through which they can develop and cultivate their talents and abilities as the former. I suspect that this is rarely the case, in fact.[2] The point is not about whether research participation offers the same kinds of opportunities for personal growth as other socially valuable undertakings. The point is to highlight the social value of each of these undertakings, the fact that each imposes risks on those who undertake them, that these risks and burdens must be freely undertaken by fully enfranchised members of a society of free and equal persons, and that conceiving of research participation on these terms entails reasonable limits on the risks and burdens to which participants can be exposed.

[2] For the claim that research participation does not offer the same goods as work, see Jonas (1969), Różyńska (2018), and Malmqvist (2019). See also Anderson and Weijer (2002), Dickert and Grady (1999), Lynch (2014), and Lemmens and Elliott (1999).

To make these ideas operational in practice, we need to understand what it takes to respect the principle of equal concern in more concrete terms. Three criteria give operational content to this requirement in practice.

6.3 Criteria for Operationalizing Equal Concern

6.3.1 No Unnecessary Risk

To facilitate concrete risk evaluations, the integrative approach defines three operational criteria to specify the terms on which important research can be advanced without compromising the status of research participants as the moral and political equals of their compatriots. Each of these criteria are to be applied to every study or set of studies under review.
The first operational criterion for ensuring equal concern prohibits arbitrary and unnecessary risks and burdens:

No Unnecessary Risk: To be consistent with the principle of equal concern, the risks to both the basic and the personal interests of participants should be reduced to those that are necessary to produce the knowledge needed to address a gap in the ability of a social system—such as a health system—to safeguard and advance the basic interests of the people they serve.

From this criterion it follows that it is never acceptable to expose research participants to risks that are gratuitous or more significant than is necessary. This requirement is not restricted to basic interests since even though the personal interests of research participants may not be widely shared, they may nevertheless be of profound importance to the particular individual. Where impositions on the personal interests of study participants are foreseeable, such impositions should be reduced so as to ensure that their presence reflects the necessity of their contribution to valuable science and not simply social disregard for their first-order life plan by researchers.

Demonstrating that those risks or burdens are necessary to facilitate the legitimate epistemic goals of scientific inquiry provides credible assurance that decisions that stakeholders make regarding the imposition of risks and burdens track the legitimate social purpose of research. In other words, it ensures that the risks and burdens to which participants are exposed cannot

be eliminated without compromising the quality or the integrity of the evidence a study is designed to generate. Satisfying this condition is necessary to ensure that the risks and burdens of research are not arbitrary impositions that reflect antipathy or indifference to individuals or their particular interests (Pettit 1997, vii).

6.3.2 Special Concern for Basic Interests

The second operational criterion for preserving equality is to be applied after the first. It reflects the special normative status of the basic interests of individuals:

Special Concern for Basic Interests: If the basic interests of research participants are threatened or impaired (for example, by sickness, injury, or disease), participants must be provided a level of care and protection for their basic interests that does not fall below what at least a reasonable minority of experts in the relevant fields (e.g., the medical or public health community) would regard as the most beneficial method of response.[3]

This requirement applies to cases in which the basic interests of individuals are threatened. This is a morally important circumstance for two reasons. First, these interests play a critical role in securing the standing of individuals as free and equal persons (§4.5). They are the rudimentary building blocks that individuals need in order to be free to formulate, pursue and revise a distinctive life plan. When the basic interests of study participants are threatened, so is the fair value of their highest-order interest in having real freedom to formulate, pursue and revise a life plan of their own. Second, a just social order is committed to securing and upholding the freedom and equality of persons. The principle of special concern for basic interests provides credible social assurance to study participants that research functions as a division of social labor in which their status as free and equal persons will be respected. As they participate in activities designed to create the means of securing the basic interests of others, study participants can be secure in the knowledge

[3] This formulation follows the practice of presupposing that the deliberations in question take place against the backdrop of a particular set of basic institutions. In chapter 9 I argue explicitly for the claim that this principle should be understood as holding relative to the level of care and protection that can be attained and sustained in the basic social institutions of the host community.

that their basic interests will not be treated with lesser regard or subject to antipathy, indifference or neglect.

If an intervention is not regarded as the best way to safeguard or advance the basic interests of a person by even a reasonable minority of experts (in other words, nobody champions that approach as among the best ways to treat, prevent, or diagnose the health problem in question) then I will refer to it as "substandard." The operational test for determining which practices, procedures, or interventions meet this standard is explicated in §6.4.

The focus here on not falling below the standard of what at least a reasonable minority of experts would recommend is meant to capture the idea that even when there is significant uncertainty or widespread disagreement about what constitutes the best or optimal response to a particular problem, it is often possible to identify interventions that would not be regarded as among those likely to be best by even a reasonable minority of the relevant expert community. When the basic interests of individuals are at stake, allocating study participants to interventions that are not regarded as among those likely to be best by even a reasonable minority of the relevant expert community violates the principle of equal concern. In such a case, although different experts might disagree about how best to meet a person's basic interests, they all agree that there are better alternatives to the intervention in question.

When the basic interests of a participant are not at stake, this requirement does not apply. This represents a major difference between the integrative approach and frameworks that first distinguish interventions that are offered with therapeutic intent from those that are employed for purely research-related purposes and then subject interventions in these two different categories to different standards of appraisal. I will return to this point in more detail in §6.7.2. For now, it is sufficient to say that the reason to focus on the interests that are at stake, rather than the rationale for providing an intervention, is that it is the interests of participants that are morally relevant. The distinction between therapeutic and purely research-related study procedures attempts to track this distinction, but indirectly.

This indirect route is problematic because interventions can be delivered with therapeutic intent even when the condition that is being treated does not implicate the basic interests of the recipient. For example, there may be circumstances where researchers want to investigate new prophylactic or therapeutic measures for minor medical conditions, such as minor scrapes and cuts or male-pattern baldness. It may be permissible in such cases to test new interventions against a baseline of no treatment, even if established

effective interventions already exist for these conditions (as long as it can be established that such a study still has meaningful social value). This would not violate the special concern for basic interests as long as the conditions under study do not threaten the basic interests of study participants.

Views that hold that interventions administered with therapeutic warrant must be tested against the best available alternative would prohibit studies that allow participants to be randomized to no treatment in such cases. In order to avoid this implication, some have argued that in such minor cases no treatment is often a legitimate therapeutic option. This observation is correct, but it only reinforces the point I am making here. The reason no treatment is a legitimate therapeutic alternative in such cases is explained by the fact that minor, transient, or cosmetic problems do not threaten the basic interests of individuals and, as such, we can legitimately ask study participants to voluntarily forego such interventions if doing so is necessary to conduct a study with the requisite social value (§6.7.2).

A just division of social labor can permit community members to ask one another to risk, sacrifice, alter, or limit ends or goals that are part of their individual life plan—their personal interests—in an effort to secure for others the basic interests that are necessary components of the freedom to formulate, pursue and revise such an individual life plan. This means that it is morally permissible for researchers to ask participants if they are willing to undergo painful but transient procedures, to endure unpleasant but temporary experiences, to bear inconveniences, or to take risks that are unlikely to compromise their basic interests but that different individuals may perceive as more or less significant depending on the way those risks relate to the constituents of their personal life plan. When the basic interests of study participants are not at risk, this permission applies to all study procedures, regardless of the warrant for their use.

When the basic interests of participants are at risk, then there is a strong moral imperative to ensure that participants receive a level of care for their basic interests that does not fall below what would be recommended by at least a reasonable minority of experts. Study participants cannot be offered a course of care for their basic interests that falls below what would be recommended by at least a reasonable minority of experts, and they cannot be subjected to study-related procedures that would compromise their basic interests. It is permissible to ask them to accept risks to their personal interests, from alterations in the course of their care or from purely research-related procedures, so long as the risks and burdens have been reduced as much as possible, are necessary for the conduct of socially valuable science,

and are consistent with the provision of care for the participant's basic interests that is not substandard.

The criterion of special concern for the basic interests of participants addresses the question of what it is permissible for researchers to offer to study participants. When research implicates the basic interests of participants, not only must unnecessary risks be reduced, but also researchers cannot offer participation in a study that would provide them with substandard care. They can, however, offer participation in studies in which it is necessary to expose participants to risks and burdens to their personal interests. In all cases this must be done in the context of a process of informed consent in which participants (or their proxies) are given a clear and accessible explanation of the rationale for such risks as well as their expected duration and magnitude. It is then up to individuals to evaluate these offers and to decide for themselves whether those particular burdens and risks are reasonable in light of the goals of the study and their personal values and commitments.[4]

6.3.3 Social Consistency

If the risks associated with research could be limited to the personal interests of participants, then the two operational criteria discussed so far would be sufficient to assess and manage research risks. The problem is that almost every activity poses some degree of risk to a person's basic interests. For example, in the vast majority of cases, a blood draw will expose most people to only brief discomfort and an unsightly blemish. Nevertheless, there is a small but non-zero probability that a blood draw could cause a fatal or debilitating infection. More invasive procedures, such as biopsies or spinal taps, may pose a higher risk of debilitating, permanent, or fatal adverse events, even though the absolute risk of these events may be quite low when they are performed by trained personnel under controlled conditions.

As a result, the integrative approach requires an additional principle for regulating the extent to which stakeholders in the research enterprise can be exposed to risks to their basic interests without violating the principle of equal concern.

[4] The wording here is not meant to imply that waivers of informed consent are never permissible. The purpose of these remarks is simply to illustrate the division of moral labor between the risk assessments that structure the way a study is designed and the subsequent requirement to seek the consent of study participants.

The integrative approach uses a third operational criterion to ensure that residual risks to the basic interests of study participants are consistent with the principle of equal concern.

Social Consistency: In all cases, the cumulative incremental risks to the basic interests of study participants that are not offset by the prospect of direct benefit to the participant must not be greater than the risks to the basic interests of individuals permitted in the context of other socially sanctioned activities that are similar in structure to the research enterprise.

The third operational criterion recognizes that respect for the moral equality of individuals cannot require that they be prohibited from voluntarily assuming risks to their basic interests. First, such a standard simply could not be achieved; even routine activities involve some incremental risk to a person's basic interests. Second, enforcing such a standard would not only rule out participating in medical research, but, as we saw in the previous chapter (§5.9–11), it would rule out pursuing a career as a researcher (and many other important social activities and professional roles), since that work can itself involve small but non-zero risks to the researcher's basic interests. The challenge, therefore, is to establish when incremental risks to the basic interests of individuals violate the underlying commitment to moral equality, recognizing that there can be reasonable diversity across a range of different social activities in the extent to which risks to the basic interests of persons can be seen as reasonable.

The third operational criterion addresses this context sensitivity by requiring stakeholders to identify social activities that are structurally similar to the research enterprise and to ensure that incremental risks to the basic interests of participants do not exceed the incremental risks to the basic interests of individuals associated with those structurally similar social activities. The central challenge then lies in delineating criteria for structural similarity that can be used to locate relevant comparison classes of activities and then in determining how to make these comparisons in practice.

The requirement of structural similarity is meant to capture the idea that it is not appropriate to use just any social activity to determine what kind of incremental risks to the basic interests of participants are morally permissible. For example, some people may enjoy auto racing, ski jumping, or hang-gliding at least partly because of the thrill that comes from their associated risks. More generally, there may be activities in which individuals willingly

engage such that eliminating all risk of physical harm would detract from the underlying value of that activity. In research, however, if it were possible to eliminate all risk of harm it should make participation in this activity more, rather than less, attractive.

This feature of clinical research should therefore be used as a criterion of structural similarity. That is, appropriate comparison classes of activities should be ones whose primary social purpose is to benefit others, where the associated risks are viewed as necessary evils such that reducing or eliminating those risks would render that activity more attractive to participants.

Additionally, when individuals perceive themselves as having control over salient features of an activity, they are often willing to accept greater risks to themselves than in similar activities where they lack such control. This may help to explain why people are willing to tolerate greater risks from driving than from airline flight or other forms of public transportation. Such asymmetries matter in the research context because research participants put their interests in the hands of identifiable parties who possess knowledge and expertise that participants lack, and who pursue a diverse set of interests, some of which may overlap or dovetail with those of participants and some of which may not. This militates in favor of comparing the risks to the basic interests of research participants that cannot be eliminated, to risks to the basic interests of community members that are associated with social activities that involve this kind of principle-agent relationship.

Finally, it is imperative to avoid using activities as comparators in which oversight mechanisms or safety regulations are poorly enforced or are widely recognized to be inadequate. For instance, coal mining has become a safer occupation then it was several decades ago because of tougher safety regulations. However, at the time that I wrote one of the papers on which this chapter is based, there were several high-profile accidents at mines that had been repeatedly cited for safety violations. As a result, the risks that coal miners face in actual practice were clearly higher than what was judged to be socially acceptable, as evidenced by the fact that actual conditions on the ground often fell short of the requirements of existing health and safety regulations. Appropriate comparator activities should not only be the subject of active public oversight, but should have a record of complying with the requirements outlined in such oversight, so that the risk profile associated with the activity can be seen, at least prima facie, as representing a level of risk that is deemed socially acceptable after due reflection.

To be clear, the goal in applying the third operational criterion is to find reasonable criteria of similarity that can be used to identify appropriate comparison classes of activities and then to examine the risk profiles associated with activities that satisfy some or all of these criteria. This process itself may require careful adjustments in the criteria of similarity as well as discerning judgments about whether the activities that meet these criteria ought to be endorsed as appropriate comparators for clinical research. This is therefore an inherently normative or evaluative process. The objective is not to avoid making such normative judgments. It is, rather, to find a reasonable set of criteria that can be used to facilitate this process so that data that exist about the risks associated with socially important activities in one sphere can be used to assess the incremental risks to the basic interests of research participants that come from purely research-related elements of a particular study.

One place to look for appropriate comparison classes of activities might be to public service professions, such as volunteer fire departments or paramedic services. The volunteer nature of these activities combined with their orientation to serving the public interest represent important structural similarities to the research enterprise. Similarly, these occupations are often subject to varying degrees of public oversight. However, because there is no principal-agent relationship in these activities it may not be appropriate to permit in clinical research activities that have a risk profile that is similar to the most dangerous activities that individuals in these roles sometimes undertake.

The idea I am proposing is to use these comparison classes of activities to construct practical tests for this third operational criterion. Such practical tests demarcate an acceptable upper bound on the incremental risks to the basic interests of participants in clinical research. Ideally, most studies would impose risks that fall well short of this upper bound. Where such risks cannot be eliminated and are necessary to produce socially valuable information, the proposal is to ensure that they are not greater than the incremental risks to basic interests that members of helping professions, such as fire fighters or paramedics, face on a routine basis. If a phase I clinical trial involving healthy participants failed to meet such a test, for example, then it would have to be redesigned, delayed until further pre-clinical research could be completed, or the data would have to be generated in some other fashion.

Finally, as we saw in §5.11.2, the limits on the risks that it is reasonable to pursue in different activities can change according to what I there called the principle of proportionality. For example, the risks to which volunteer paramedics or fire fighters can be exposed in the course of protecting property, or safeguarding individuals from threats that are more distant in time, is often lower than the risks to which such persons can be exposed when the risks to others are greater and more immediate. So too, then, permissible research risks can be subject to a similar requirement of proportionality. When research is in early stages, the permissible level of residual risk to the basic interests of study participants should be lower than when those risks are necessary to ensure that results are more directly applicable to patient care or clinical practice.

6.4 Uncertainty as a Practical Test within the Integrative Approach

6.4.1 Uncertainty Regarding Basic Interests

Because the integrative approach rejects the two dogmas of research ethics discussed in the last chapter, it can fill out the template for the appeal to uncertainty (§5.4) in a way that is both conceptually and practically coherent. To see this, we must first articulate the practical test that is to be used to determine whether or not a particular clinical trial satisfies the second operational criterion outlined previously—special concern for basic interests.

To construct a practical test for the second operational criterion we need to know when it is ethically permissible to allocate a study participant to a given intervention and when doing so constitutes substandard care and therefore represents a violation of equal concern. The integrative approach uses the following definition of admissibility to construct this practical test.[5]

[5] More formally, let $I = \{1,...i,...n\}$ be the set of individuals with a particular medical condition for which there is a set of available treatment options $S = \{s_1...,s_m\}$. Let U_i be the set of interventions from S to which individual i might be allocated within a particular clinical trial and let U_i^* be the set of interventions from S that are admissible treatment options for the individual i.

Uncertainty Regarding Basic Interests: A treatment s_j is admissible for individual i just in case there is either uncertainty among, or conflict between, expert clinicians about whether s_j is dominated by any other members of S as a treatment for individual i. For each individual in I, the care and protection afforded to that individual's basic interests falls within the threshold of competent medical care just in case each intervention in U_i is a member of U_i^*.

Uncertainty Regarding Basic Interests: For each individual with a particular condition (e.g., a health problem), the care and protection afforded to that individual's basic interests satisfies the condition of equal concern just in case every intervention to which that person might be allocated in a research study is admissible. An intervention is admissible for an individual just in case there is either uncertainty among, or conflict between, experts about whether it is dominated by any other intervention as a means of safeguarding the basic interests of that individual.

Notice that this practical test is formulated at the level of the individual study participant. This addresses a concern raised by Miller and Weijer (2006b) and discussed in §5.7.5, namely, that although medical experts might be uncertain or disagree about the merits of a particular intervention for patients with a particular medical condition, they may not be uncertain about its merits for any particular individual. This might happen, for instance, if that individual has a medical condition that is clearly contraindicated, putting them at elevated risk were they to receive a particular treatment.

This practical test requires that each potential study participant can only be invited to participate in studies that allocate them to admissible interventions. In this respect, it is similar to the criteria of admissibility recommended by Kadane and colleagues (1996). In that trial, a treatment was deemed to be admissible for a particular participant just in case it was judged to be the best treatment option for that individual by at least one from among a set of expert clinicians. In this case, however, the expert clinicians were actually computer models that had been constructed out of a careful elicitation process involving real clinicians. A less computationally complex solution to this problem is for each participant to be screened by several experts. If different experts who each prefer one intervention over another for a particular medical condition regard their favored intervention as admissible for this particular participant, then it is permissible to allow that individual to be randomized to those interventions.

The most salient difference between the concept of admissibility defined here and the one articulated by Kadane and colleagues is that the integrative approach limits the scope of these judgments to the basic interests of study participants. Nevertheless, when this condition is met, each individual who participates in a clinical trial is assured of receiving a package of medical

care that would be recommended for them by at least a reasonable minority of expert clinicians.

This practical test is also similar to Freedman's clinical equipoise, but here too there are some important differences. First, the moral force of this requirement is grounded, not in the individual physician's therapeutic obligation, but in the fundamental importance of the basic interests of individuals as both a target for the research enterprise and as a constraint on the way it is organized and operates. The judgments of particular experts, such as medical professionals when evaluating medical research, are used only to determine the practices, procedures, or interventions that represent the most effective means of safeguarding the basic interests of the individual in question. They do not play any role in grounding the normative foundation of the integrative approach.

Second, this practical test explicitly distinguishes agnosticism from a state of clinical conflict between experts (§5.6.3). The former obtains when individual clinicians have no ground for preferring one treatment from among the set of available options over any others as a treatment for that individual. This state might occur, for example, when a novel intervention begins to show sufficient promise in animal models and in early trials in humans that clinicians become uncertain about its net therapeutic advantage relative to existing interventions for some set of individuals. When this occurs, it may be permissible to initiate a clinical trial in which individuals for whom both of these interventions are admissible are randomized to one of them.

Clinical conflict exists when individual expert clinicians have definitive expert assessments that one intervention is preferable to the other options for a particular individual, but different experts prefer different interventions from this set of options for that individual. So, for example, one expert might regard high-dose chemotherapy with autologous bone marrow transplant (HDC-ABMT) for end-stage breast cancer as the best treatment option for a particular patient. A different expert may regard continuation of standard chemotherapy as preferable to this more aggressive intervention for the same patient. In this case, offering this person the option of participating in a clinical trial in which she might be randomized to either of these treatment options is ethically permissible since, no matter what the result of the randomization, this person is guaranteed to receive an intervention that would be recommended for her by at least a reasonable minority of expert clinicians.

6.4.2 Reconciling Social Value, Concern for Welfare, and Equal Concern: HDC-ABMT as an Example

The integrative approach provides a clear rationale for studies like the landmark trial conducted by Stadtmauer and colleagues (2000) that tested HDC-ABMT against standard doses of chemotherapy in patients who responded well to an initial 6–8 week course of induction chemotherapy. While some clinicians were passionate proponents of HDC-ABMT, others were skeptical that its benefits outweighed its significant burdens. In some cases, the degree of polarization may have been sufficiently high that members of the various camps regarded those who did not share their treatment preferences as violating the clinical judgment principle discussed in §5.7. Nevertheless, this study helped put to rest a decade-long debate about the relative clinical merits of HDC-ABMT, showing that it offered no advantage over standard of care (Mello and Brennan 2001).

For women in this trial, both HDC-ABMT and conventional doses of chemotherapy were admissible treatments because there was no consensus in the expert community that either one of these interventions dominated the other. In fact, the informed expert clinical community likely divided into two camps. The first would have recommended the more aggressive treatment to their patients. The second would have recommended conventional treatment to their patients. There might have been a third camp who, seeing this polarization, was uncertain about the relative merits of these therapeutic alternatives. But the integrative approach does not presuppose or require the existence of such a third group to justify conducting a well-designed clinical trial.

Without a randomized clinical trial, an informed and well-resourced patient with end-stage breast cancer might have sought a second opinion. Had they encountered clinicians from each of these camps, they would have been told that HDC-ABMT is their best option and also that it is not their best option. They would also have been told that conventional chemotherapy is their best option and also that it is not their best option. Faced with these conflicting recommendations, such a patient could arbitrarily decide to accept the recommendation of one of these groups or could have decided to flip a coin.

The opportunity to participate in a randomized, controlled, clinical trial presents patients in this situation with the opportunity to receive a modality of care that would be recommended for them by at least a reasonable minority of experts, but under conditions that facilitate valid inference about

the therapeutic effects of these alternative interventions. In this respect, the clinical trial offers patients an opportunity to contribute to the common good without knowingly sacrificing their basic interests in the process.

Because this study satisfies the conditions outlined here it has a strong claim to satisfying the social value requirement. The knowledge it was designed to produce is necessary to improve the ability of health systems to safeguard and advance the basic interests of people with breast cancer. A trial designed to resolve conflict or uncertainty about how best to manage this fatal medical condition generates information that bridges a knowledge gap concerning how best to effectively, efficiently and equitably address an important health need of this community.

Trials that satisfy the conditions of the integrative approach also satisfy the principle of concern for welfare (§5.4 and §6.2.1). In particular, if it was consistent with concern for welfare for informed and conscientious clinicians to prescribe continued chemotherapy for a patient—to ensure that the patient received that modality of care with certainty—and for other informed and conscientious clinicians to prescribe HDC-ABMT, then it cannot violate concern for welfare to allow that same patient to be randomized to these interventions. Ensuring that no individual in the study is allocated to an intervention that is substandard, as defined here (§6.3.2), thus ensures that studies with a strong prima facie claim to producing socially valuable information are consistent with the requirement of concern for welfare.

Studies that satisfy these conditions are also consistent with the principle of equal concern. First, the risks to which participants are exposed have been reduced to those that are necessary to answer an important medical question (the first operational criterion). So, no participants are asked to bear a burden or to be exposed to a risk that is gratuitous or unnecessary. Second, the level of care and protection for the basic interests of each person in the study does not fall below what at least a reasonable minority of the expert medical or public health community would regard as the most beneficial method of response (the second operational criterion). So, no individual in the study is knowingly subjected to substandard care. Third, the residual risks to participants that are not offset by the prospect of direct benefit to a participant's basic interests are not greater than what it would be permissible for individuals to undertake in the course of a comparable social activity. In this respect, individuals who participate in research that meets these requirements are not treated with less respect or concern than the population of persons who do not participate in the study.

As a result, individuals could participate in this study secure in the knowledge that their basic interests would be respected and that their status as

free and equal persons would not be denigrated in the course of study participation. Although which modality of care they receive is determined by a random process, the care that they in fact receive is not worse than what would be recommended for a person who opted not to participate in the study. In that sense, participants are not subject to a standard of care that is worse than what is available to other study participants or to patients outside of the trial.[6] Individuals who did not want to take up, as a personal goal, the project of determining the relative clinical merits of these interventions were free to refuse to participate. For those who saw answering this question as a worthwhile project to adopt, participation offered an avenue to advance the common good without compromising their status as free and equal.

What about the charge that as data from this study accrue, the states of uncertainty or conflict that justify its continuation are likely to evaporate long before the trial reaches statistical significance (Marquis 1983; Gifford 1986; Hellman 2002)? In the previous chapter I argued that this was a powerful objection against all views that locate the relevant uncertainty in the mind of the individual clinician. I also argued that although Freedman took the critically important step of moving the relevant uncertainty out of the mind of the individual clinician and into the larger expert medical community, his approach suffered from the fact that its normative foundations were still located in the moral obligations of medical professionals. For this reason, I want to give a short answer to this question here that I will then unpack and elaborate in more detail in the next section.

It is a strength of the integrative approach that the criteria for admissibility articulated in the practical test for the second operational criterion are tightly connected to the social value requirement. When there is uncertainty or conflict among experts about how best to safeguard or advance the basic interests of individuals, trials that are designed to eliminate this conflict or uncertainty have a strong prima facie claim to generating social value. Imagine now that an interim analysis of data was pre-planned. If the evidence generated by such a trial at this interim analysis is sufficient to create consensus about the superiority of one option over the other, then the trial will

[6] It is important to note that the equipoise requirement is often charged with myopically comparing the relative merits of interventions on offer within a clinical trial without taking note of the care that might be available outside of such a study (see Kukla 2007). So it is important to emphasize that the position outlined here avoids this problem by requiring that study participants not be treated worse than other study participants and other members of the community whose basic social structures the research is designed to strengthen. See §2.6.3 note 12 and chapter 9 for further discussion.

have served its social function and ought to be terminated. In other words, if the evidence is sufficiently compelling that conscientious and fully informed experts shift their expert medical opinion in favor of one option and against the rest, then the study has served its legitimate social purpose and should be terminated.

If, in contrast, only some fully informed and conscientious clinicians are convinced and a reasonable minority remain uncertain or would continue to make clinical recommendations that conflict with those of their peers, then we have moved from either a state of uncertainty to one of clinical conflict or we have reduced but not eliminated the conflict in informed and conscientious medical judgment. In that case, it is permissible to continue the trial since both of the interventions in question remain admissible.

When interim evidence is sufficient to alter the judgment of some clinicians, that alone is not sufficient to warrant stopping the trial. The question is whether, if the trial were stopped, other conscientious and informed experts would continue to recommend the different treatments in the trial to their patients. If so, then the conflict in expert medical judgment, and with this the diversity in actual treatment practices, would persist. As a result, stopping the trial makes no person better off, but it detracts from the social value of the study. Continuing the study ensures that the trial satisfies the social value requirement without requiring any individual to be allocated to substandard care.

Another way to state this is to say that the close connection between the social value requirement and the criteria for admissibility in the integrative approach help to ensure both that trials continue until they serve their legitimate social function and that the basic interests of participants receive the same degree of care and concern as is shown for individuals outside of research.

6.5 A Social Model of Learning, Uncertainty, and Disagreement

6.5.1 The Most Difficult Case: Response Adaptive Randomization

It is a theme of this book that some of the shortcomings of orthodox research ethics stem from the way it frames the core issues of the field as

being situated within the relationship between individual researchers and individual subjects. The two dogmas of research ethics embody and reinforce this framing and, in doing so, they exaggerate the extent to which tensions within the research enterprise are taken to constitute moral dilemmas that require empowering some entity within this system to make tragic choices.

The integrative approach connects conflict and uncertainty in the relevant expert community with a normative standard for regulating research that is also social in nature. This symmetry allows for a closer connection between uncertainty and the social value requirement, on the one hand, and conflicting expert judgment and concern for welfare, on the other. We can illustrate the advantages of this connection by turning to what is likely the most difficult case for approaches to risk management that appeal to uncertainty, namely, studies that incorporate response adaptive randomization (RAR).

In most clinical trials, a decision is made at the time a study is designed about the proportion of participants who will be allocated to a novel intervention and the proportion who will be allocated to the comparator (which might be the current standard of care or a placebo control). This allocation is usually fixed, in the sense that it does not change throughout the course of the trial. If the chosen allocation is 50:50, then a fair coin (or its computational analogue) is flipped to decide which intervention to provide to each participant. More generally, trials with an equal allocation assign each participant a $1/n$ chance of receiving each of the n interventions on offer in a trial. In some cases, this allocation is fixed but unequal. So, for example, a trial might have a fixed 60:40 allocation in favor of the investigational intervention.

Generally, when trials involve a fixed randomization allocation (FRA), their sample size is calculated at the time the study is designed and then a specified number of participants are recruited and allocated to the interventions in the trial's various arms. It is common for data and safety monitoring boards (DSMBs) to take an "interim look" at the data to make sure that any effects observed in the trial to that point are not so dramatic that the trial should be stopped, either because some intervention is performing extremely well or extremely badly. These interim analyses have to be planned in advance and the power of the study to detect effects of a given size has to be adjusted accordingly.

Trials that use RAR differ from this model in several ways (Lin et al., 2016). In particular, a study might begin with an equal probability ($1/n$) of participants being assigned to each of the n interventions in its various arms, but after a predetermined number of outcomes are observed (a "block" of participants), the randomization allocation is changed. The proportion of participants allocated to the various arms in a trial is altered dynamically depending on the outcomes that are observed from blocks of patients. If more promising outcomes are observed in the arm in which participants receive intervention A, for example, then the probability that participants in the next block will be allocated to A is increased. If the trial has three arms, then the probability that they will be allocated to arms B or C will be lower than the probability of being allocated to A. The relative likelihood of being allocated to B or C might remain equal or it might also be altered in favor of one of those interventions (e.g., B) if it performs worse than A but better than C. The rate at which these proportions change is determined by a function that is pre-specified at the time the study is designed.

How the randomization allocation in a study using RAR changes will depend, in part, on the outcomes that are observed in each of its arms. One of the advantages of this kind of design is that interventions that do not perform well will have their allocation proportion reduced as the trial proceeds. Normally, some threshold will be specified in advance for dropping underperforming interventions from the study. For example, if its fraction drops below 10%, then it might be dropped from the study.

RAR is often an element of study designs that allow new interventions to be added to ongoing trials without having to design a new protocol to test this new intervention (Berry 2011; Lewis 2016; Saville and Berry 2016; Trusheim et al. 2016; Renfro and Sargent 2017; Angus et al. 2019). The flexibility of being able to drop or add arms within the same study protocol makes such designs particularly attractive to a variety of stakeholders as it offers the prospect of reducing delays associated with designing and receiving approval for new studies. For example, this is attractive for pharmaceutical companies because reducing delay can increase profits by increasing the amount of time the firm has exclusive right to sell a drug. It is also attractive to patients to the extent that shorter research timelines mean that new interventions might be available sooner rather than later.

6.5.2 The Virtue of Rational Expectation

Proponents of designs that incorporate RAR argue that they are more attractive for participants because dynamically updating the randomization allocation in light of observed outcomes allows trials to allocate more participants to interventions that are performing well (Meurer, Lewis, and Berry 2012; Lewis 2016)[7]. This increases the probability that a participant in a trial with RAR will receive a direct medical benefit from participating, relative to the probability of receiving such a benefit in a trial with a FRA. In effect, the proponents of RAR argue that it is superior to designs that use FRA on the following ground.

Rational Expectation: If in expectation a participant has a greater probability of being allocated to what turns out to be a superior intervention in study design F than in design G, it is rational for that participant to prefer design F to G.

6.5.3 Does Rational Expectation Violate Concern for Welfare?

Critics argue that trials of this kind reveal a fundamental moral dilemma for research ethics because any design using RAR that satisfies rational expectation must violate concern for welfare and, with this, the principle of equal concern. The rationale for this claim is stated by Saxman (2015), although not quite in these terms. Phrased in the terms I am using here, the charge is that when the randomization weights are $1/n$ the study might respect concern for welfare and the principle of equal concern. But once evidence emerges that one intervention (e.g., A) produces better outcomes than the others, and randomization weights adjust in favor of A, then it violates concern for welfare to allow subsequent patients to be randomized to B or C.

[7] "Advocacy of adaptive designs is predicated on the belief that such novel designs will result in fewer numbers of subjects having to participate and receive an 'inferior' treatment during the research process" (Laage et al. 2017, 192).

This objection can be pressed further. If we assume that studies involving RAR satisfy the principle of rational expectation, then critics of this design can be seen as worrying that it purchases whatever gains in efficiency it offers, at the cost of violating the principle of equal concern. In particular, although more participants will be allocated to what turns out to be the most beneficial intervention (if there is one) in the trial, it is nevertheless the case, they claim, that a smaller proportion of study participants are knowingly allocated to arms of the trial that are believed to produce outcomes that are inferior to those that could be produced if they were allocated to A.

6.5.4 No Impermissible Gambles

Additionally, RAR faces a criticism that is often lodged against unequal randomization more generally, namely, that even in the best case such designs reduce the burdens on some participants by knowingly exposing other participants to care that is credibly believed to be inferior. But if it is wrong to allocate someone to intervention C with certainty, when A is available—to just give them C instead of A—then we cannot make it permissible to give that person C when A is available by reducing the probability that they will be given C. Doing this violates the following principle:

No Impermissible Gambles: If it is impermissible to directly give intervention C to a person (give it to them with probability 1) when some other intervention A is available, then it is impermissible to include C as an option in a design that would randomize that person to C with any positive probability when A is available.

The point of the exposition so far is not to defend the methodological merits of study designs that use RAR, but to formulate what looks like an extreme example of the objection we saw in the previous section. In other words, if we assume that proponents of RAR are correct when they say that it offers a range of methodological advantages over studies that use an FRA, then we appear to find ourselves immediately back in the jaws of an in-principle moral dilemma: if the study satisfies rational expectation then it appears to violate the principles of equal concern, concern for welfare, and no impermissible gambles.

6.5.5 Forcing Uncertainty into the Model of a Single Decision-Maker

Although the argument that RAR produces a conflict between the principles of rational expectation, no impermissible gambles, concern for welfare, and equal concern sounds like a novel objection, it hinges on a criterion for admissibility that is derived from the judgment of a single expert. The difference is that, in this case, the trial itself is regarded as the single expert whose views are used to represent the relevant uncertainty. It is useful, therefore, to explain why this perspective seems so appealing and what is wrong with it. In the next section I then show how RAR can be thought of as helping us to model a moderately idealized learning health system and that, when studies are designed on these terms, the apparent dilemma disappears.

The reason that studies that use RAR seem to pose a deeper problem for approaches to research risk that appeal to uncertainty is that it encourages the idea that we can regard the randomization weights in a trial as expressing the preference of the medical community for the best performing intervention. When the randomization weights are $1/n$, for each of the n interventions, then the trial appears to be indifferent between those n interventions. But once the weights shift in favor of one intervention, the trial is no longer indifferent. It favors one option over the others. In effect, the trial is treated as a meta-agent constructed by taking a weighted average of the opinions of the different treatment communities. In that sense, the critique of RAR seems novel and interesting because this meta-agent that the trial is treated as modeling is easily seen as occupying a privileged epistemic state such that its judgments ought to be normative for the practice community in a way that the judgments of ordinary individuals might not.[8]

This view is alluring because it appeals to the idea that rational inquiry requires an agent, it treats the trial as such an agent, and it treats the randomization weights as though they are that agent's degrees of belief about the relative merits of the interventions in question. This social agent is created by aggregating the judgments of the diverse experts in the community, combining them into a higher-order decision model. In essence, it assigns a weight to the likelihood that each expert is correct and then chooses in a way that maximizes expected value. This is a concrete example, in microcosm, of a larger view of scientific consensus that many find intuitively appealing,

[8] Leonard Savage attributed a similar view to Woodbury (Savage 1972).

namely, that the goal of scientific consensus is to take in the diversity of beliefs in the scientific community, assign them weights, and form a single all-things-considered model out of this diversity.

This view faces several problems. First, it is a version of the linear opinion pooling rule for combining individual judgments into a group or social judgment. But the social agent constructed by assigning weights to the views of the individual experts can make recommendations that radically diverge from the recommendations of all of the experts from which it is created. For example, each expert may regard certain events (the temperature in Beijing today and whether to use treatment A or B for a certain patient in New York) as probabilistically independent and, as a result, would not base treatment decisions on what he or she recognizes as an irrelevant event (no medical expert will decide the merits among rival treatments for a patient who resides in New York by asking what the weather is that day in Beijing). But these relationships of probabilistic independence are not generally preserved in the linear opinion pool. As a result, the "social agent" can change its treatment recommendations upon learning the weather in Beijing, even though no particular expert would do so.[9] It is not clear why any expert should see as normative a model that would change its treatment preference on the basis of information that is not regarded as relevant by any of the models from which it is constructed.

Second, we already assumed from the beginning that every expert knows that many other, equally well-credentialed and informed experts make treatment recommendations that conflict with their own. So, it is not clear why adding another expert to the mix (in this case, the expert created by aggregating the views of the community) should alter the fact that the original experts do not find the existing evidence sufficiently persuasive to alter their treatment practice.

Third, there is the problem that, given the impoverished nature of our understanding of the underlying causal structure of health problems, experts

[9] Genest and colleagues (1986) establish that in order to be "Externally Bayesian" the pooling rule has to be of the "logarithmic" form. For the purposes of the present argument, it is sufficient to see that one aspect of the Externally Bayesian condition is that when experts regard two events (A and B) as independent, this condition requires preserving the expert judgments that each event is irrelevant to the probability for the other ($P(A \mid B) = P(A)$), after pooling. Treating a trial as a social agent involves creating a social consensus model by taking weighted averages of different treatment communities. But weighted averages are a convex combination and, as Seidenfeld and colleagues (2010) show, a convex combination of expert judgments is not a logarithmic rule. As a result, the social agent discussed above cannot preserve the judgments of experts about which events are relevant to treatment decisions.

have a difficult time predicting which theories of disease or interventions are likely to be correct or best (Kimmelman and London 2011). In this environment there are significant dangers from group-think, the situation in which experts too quickly converge to the same view of a problem. If experts in a community too readily update their beliefs on the basis of what their colleagues regard as persuasive, then spurious results that are bound to happen as a matter of chance can cause such communities to prematurely adopt promising but ultimately false beliefs.

Communities with more diversity among experts are more productive in the sense that they are less prone to converging on false answers and more efficient at exploring alternatives. As a result, communities with this kind of diversity are better at locating effective solutions to pressing problems (Hull 1988; Solomon 1992; Kitcher 1995; Zollman 2010; Muldoon 2013).

The point for our present purposes is that we must be wary of research methods that impose consensus on conscientious and informed medical experts in cases where the available medical evidence is not sufficient to alter their medical practice. Experts who embrace different theories of disease pathology and intervention mechanism are likely to update their beliefs at different rates in the face of the same evidence. Reasonable, transient diversity among experts is not simply a descriptive feature of many actual scientific communities, but a normatively desirable feature that plays an important epistemological role in the health and fecundity of those communities (Zollman 2010). In light of this, trials must be designed with rigorous methods so that the evidence they produce is viewed as credible by reasonable experts. Although such experts may change their beliefs at different rates, the goal is to produce the evidence that these stakeholders need to alter their assessments of interventions or strategies and, ultimately, to improve clinical practice.

6.6 Modeling a Learning Health System

6.6.1 Reasonable Diversity of Conscientious and Informed Experts

It is important to recognize that RAR can be a useful part of a learning health system if it is an element of trials designed to model the transient diversity of reasonable experts without amalgamating their beliefs into a single,

meta-model of uncertainty. Demonstrating how this can be accomplished makes it easier to see how the objection discussed in the previous section reflects the dogmas of research ethics that I have argued should be rejected. It also illustrates how the integrative approach provides concrete guidance for ensuring that studies can meet the social value requirement while continuing to adhere to the principles of respect for welfare, no impermissible gambles, and equal concern.

If studies are to provide evidence that will change clinical practice, then they must be designed in ways that reflect and address the reasonable diversity of expert opinion within communities of informed and conscientious experts. Consider again a case in which there are three interventions, A, B, and C for treating patients with a particular illness and that at least a reasonable minority of experts favor A, as a treatment for this condition, while other experts favor B, and still others favor C. According to the integrative approach, if experts from each of these groups would recommend their favored intervention for a particular patient (providing it to that patient directly so that there is no uncertainty about which intervention the patient receives), then it cannot violate the principles of respect for welfare or no impermissible gambles to allow that same person to be allocated to one of these interventions by a random process.

Imagine that names of the experts from each of these groups were placed into a hat. Individuals draw the name of an expert from the hat and that expert evaluates them and then recommends their favored treatment unless there are specific reasons to avoid this intervention for the person in question. In this situation, the probability that a patient would be treated by an expert from an A-favoring community would depend on the number of A-favoring experts in the bag, relative to B- and C-favoring experts.

Imagine further that after a name is drawn and the recommendation given, the name is returned to the bag. After a block of results are observed, every expert whose name is in the bag updates their beliefs. If the block of observed results favors A, then experts who favor A will favor it more strongly. Some experts who weakly favored B or C may now favor A, while others become uncertain about the relative merits of A versus their previously favored interventions. Some who more strongly favored B or C may continue to favor these interventions but less strongly than before. This process is repeated. If, at some point, the evidence against one intervention, C, is so consistent that the number of C-favoring experts drops below a specified threshold, then we might remove C-favoring names from the bag because we judge that they no

longer represent a reasonable minority of experts. The key point is that as evidence emerges and the beliefs of these individual experts change, the relative size of the communities that favor each intervention will grow or shrink.

This model provides a framework for designing trials that use RAR so that they model a health system that is moderately idealized in this respect: the beliefs of the experts in these communities reflect the beliefs of fully informed and conscientious experts in the real world, with the idealization that when evidence emerges about the relative merits of a set of interventions, the experts update their beliefs about the relative merits of those interventions on the basis of that information. Because these experts reflect the diversity of reasonable and informed expert opinion in the real world, they should agree about when evidence favors one intervention over the others, but they may differ in how they respond to this information. If evidence that favors A emerges, experts who favored A will now favor it more strongly. Some experts who were uncertain may now favor A, but others may remain uncertain. Some who favored B or C may now be uncertain, but others may continue to favor B or C, only slightly less strongly.

On the model I am proposing, randomization weights in a study that employs RAR do not reflect the beliefs of any agent. Instead, they should reflect the relative proportion of experts who, in a moderately idealized community of experts, would recommend each intervention in the study. This approach captures the idea that if a trial were stopped after only the initial block of evidence, some experts in the actual medical community would continue to favor and to recommend B or C. As long as the community of experts who favor B or C constitutes at least a reasonable minority, then it remains permissible to randomize patients to these trial arms. It remains permissible because participants within such a trial are not provided with substandard care, as defined in §6.3.2—they continue to receive a level of care that would be recommended for them by at least a reasonable minority of expert clinicians.

The goal is for the trial to capture the extent of the uncertainty about, or conflict over, the relative merits of interventions in a way that reflects the diversity of real experts while ensuring that these experts update their judgments in light of emerging evidence. In this way, diversity should be reasonable—experts change their judgments in light of emerging evidence, although different experts may change their judgments at different rates. This is meant to exclude situations in which diversity persists because experts are unaware of evidence, because communities are wedded to tradition rather

than scientific information, or because communities are in some other way impervious to evidence.

6.6.2 Reconciling Social Value, Concern for Welfare, Equal Concern, and No Impermissible Gambles

When studies model the beliefs of conscientious and informed experts and are designed to generate the evidence that such experts are likely to regard as credible, then they have a strong prima facie claim to social value. The reason is that they are designed to generate evidence that is likely to alter the practice of experts in the relevant medical community. Thinking of trials on the model I've described in this section, and designing them accordingly, helps to ensure that studies with social value continue until they generate the evidence necessary to alter practice in the expert clinical community.

Altering the randomization weights in this framework does not violate concern for welfare. From the fact that some particular expert is persuaded that A produces better outcomes than B or C it does not follow that all other conscientious and well-informed experts who see this same information will find it compelling enough to shift their treatment recommendation to favor A. If it is permissible for a patient to be treated in clinical practice by practitioners from clinical communities that each favor one of these interventions, then it follows that A, B, and C are all admissible treatment options for that patient. If it is consistent with concern for welfare for a patient to be directly treated with A or B or C (to receive that intervention with certainty from a conscientious and informed expert who regards it as best for the person in question), then it cannot violate concern for welfare if that patient is assigned to those interventions with any distribution of probabilities that sums to 1. Even if every clinician in these treatment communities has a strict preference over the available treatment options (nobody thinks the probability of success for each is 1/3), the condition of uncertainty over basic interests exists between these treatment options, and no set of randomization weights that sums to unity is impermissible. As long as the admissibility criterion outlined in §6.4.1 is satisfied, continued randomization is consistent with the principle of concern for welfare.

Altering the randomization weights in this framework does not violate the principle of no impermissible gambles. The criterion of admissibility defined here prohibits participants from being randomized to interventions

that are substandard in the sense that they would not be recommended for that person by even a reasonable minority of experts. As a result, all of the gambles permitted by the integrative approach are morally permissible.

The arguments made here and in the previous section demonstrate that trials employing RAR can be designed on terms that are consistent with the principle of equal concern as long as they satisfy the conditions of the integrative approach. Stopping such a study when some particular expert or group of experts find its results persuasive is thus self-defeating if at least a reasonable minority of other conscientious and informed experts who look at the same information would continue to provide B or C to their patients.

6.6.3 More General Relevance of the Result

The model proposed here for dealing with RAR can be used to think about randomized clinical trials that employ a fixed randomization design as well. The only difference is when experts see the data on which they update their beliefs. In other words, we can think of a traditional FRA design as one in which a participant draws the name of an expert from the hat, the expert makes a treatment recommendation, and then the name is returned to the hat. This process continues until a predetermined number of patients have been treated by experts from the different groups. The predetermined number should be calculated on the basis of the strength of evidence that will be needed to alter practice in this moderately idealized learning health system. Once that number of participants has been treated, the outcomes are revealed and the experts that make up these treatment communities alter their beliefs. Rather than updating their beliefs on outcomes as they emerge in the study, experts update their beliefs once all participants have received the relevant interventions.

Here again, one key point is that such a study must be designed to detect effects that experts regard as meaningful, with sufficient power that the results of the trial can credibly be expected to change the practice of the experts in these communities. The main difference in studies that use RAR is that individual experts are given the opportunity to change their beliefs on the basis of evidence as it emerges from the trial. The proponents of RAR hold that, if properly designed, this form of adaptation is morally superior because it satisfies the principle of rational expectation. On average, fewer participants will be allocated to study arms that are ineffective or harmful

than are allocated to such arms in a design that employs a fixed randomization scheme.

For our present purposes, the key point is that the use of RAR does not conflict with the key ethical principles I have outlined here. The model I have described in this section is useful because it illustrates how a social understanding of conflict and uncertainty avoids the errors of appeals of uncertainty that focus solely on the beliefs of a single agent. It is also useful for the way it highlights the connection between the criteria for admissibility and the transient diversity of expert beliefs in order to ensure that studies fulfill their social purpose and shift the practices of the informed and conscientious experts on whom community members rely to safeguard their basic interests.

6.6.4 The Limit of Reasonable Diversity

At this point it might be objected that the approach I have described here requires an account of when we should stop regarding a minority of the medical community as reasonable and view their treatment preferences as no longer a part of the standard of care. This is indeed an important and pressing problem. But it is one that we face whether or not we embrace the integrative approach. After all, the moral purpose of medical research is to alter clinical practice in ways that improve the ability of health systems to meet the basic interests of community members. When experts are not conscientious, when they do not continue their medical education or do not update their beliefs on the basis of evidence, then that is a problem for any approach to research and research ethics. Alternatively, when conscientious and informed experts disagree, we must resist trying to settle those disagreements by fiat. In some cases, minority opinions turn out to be correct and the received wisdom is wrong. The best practice is to use well-designed studies to generate the evidence that conscientious and informed experts use to change their beliefs and to make sure that studies are designed to generate that evidence.

It is an advantage of the approach I describe here that it highlights the importance of ensuring that the beliefs of the idealized communities reflected in any study design capture the enthusiasm of real-world clinician-researchers for the various interventions for a medical condition as well as the more conservative or skeptical views of other experts. Explicit decisions can then

be incorporated into the trial about when a community's views should be regarded as no longer reflecting the practice of responsible medicine.

In other words, the integrative approach helps to align the design of trials with their social purpose of improving the capacity of health systems to effectively, efficiently, and equitably meet the needs of those they serve. It also ensures symmetry between the practices that are regarded as ethically permissible outside of a trial and the practices that are permissible within the trial. The key point is that the integrative approach preserves these symmetries—participants in clinical trials are not deprived of a level of care that would be recommended for them by at least a reasonable minority of experts, and studies that are designed to generate the evidence that such experts need to alter their practice have a strong, prima facie claim to generating socially valuable information.

6.7 The Integrative Approach versus Alternatives

6.7.1 Criteria for Evaluating Alternative Frameworks for Risk Assessment and Management

Up to this point the goal of the present chapter has been to articulate the key elements of the integrative approach to risk assessment and management and to show how it reconciles a set of ethical requirements that capture ethical concerns that have traditionally been treated as incompatible or irreconcilable. The rest of this chapter is aimed at clarifying the merits of this approach relative to the main alternatives.

When comparing the merits of alternative frameworks for evaluating and managing risk in research, three broad criteria are relevant. The first is their *normative foundation*: an acceptable framework for risk assessment should ground its key insights and requirements in values that stakeholders can recognize as legitimate for guiding and restricting their conduct in this domain. This justification should also provide a coherent standpoint of sufficient generality that it applies to the full range of cases that occur in research involving humans. If all else is equal, frameworks that achieve a greater range of applicability without recourse to ad hoc, ancillary principles should be preferred to frameworks that require the addition of such principles.

The second criterion is the *appropriate integration* or reconciliation of the distinct concerns to which any such framework must be responsive. As we

saw in the previous chapter, in diverse communities where people are free to cultivate a wide range of life plans, different individuals are likely to have attitudes about the reasonableness of various risks that disagree or conflict with the attitudes of others. While some may be particularly averse to certain types of risk, others may be willing to accept quite significant personal risk for the prospect of advancing socially meritorious projects. An acceptable framework will need to promote socially valuable research, demonstrate respect for individual autonomy, and recognize the social value of undertaking risk in the pursuit of socially valuable ends, while providing credible safeguards to protect the rights and welfare of study participants. To the extent that this involves setting limits on the risks to which participants may permissibly be exposed in research, an acceptable framework should set such limits on the basis of reasons that all stakeholders can recognize as nonarbitrary restrictions on the pursuit of their own ends and projects.

A third criterion is *operational clarity*. As a branch of practical ethics, the guiding ambition of research ethics is to provide a framework for evaluating clinical research that can help stakeholders resolve reasonable disagreements in a way that is publicly accessible and defensible. As such, an acceptable framework should not only ground its requirements in a unified normative perspective, but it should (1) elucidate a set of operational criteria or markers that delineate the parameters or boundaries that separate reasonable from excessive risks and (2) articulate practical tests that deliberators can use in order to determine whether or not these operational criteria have been met in any particular case.

It is my contention that the integrative approach fares better on these criteria than the available alternatives. In §6.7.2 I address other approaches that, like the integrative approach, attempt to ground their framework in an appeal to uncertainty. In 6.7.3 I address approaches that reject an appeal to uncertainty and instead adopt a more consequentialist approach.

6.7.2 Other Appeals to Uncertainty and Component Analysis

It is a strength of the integrative approach that it is grounded in a unified moral and political foundation that is normative for anyone motivated to respect other persons as free and equal. In particular, in diverse communities, different individuals will develop different life plans within which various

activities take on particular personal importance or significance. As a result, individuals in such communities may disagree about a range of issues, including the value of various life plans, the value of various risks in relationship to particular ends and the place of health and health-related values in their individual and shared conceptions of the good life. To ensure that the activities of social institutions, including those that regulate the conduct of research, are not asserting arbitrary social authority, the members of such communities require a social standpoint from which they can evaluate the extent to which both impositions of risk and limitations on permissible risk are socially justified and not morally arbitrary. This standpoint need not be one that individuals embrace as part of their first-order conception of the good or the good life, but it must be a standpoint they are capable of inhabiting and which they can recognize as morally authoritative for regulating social institutions.

The integrative approach constructs the required standpoint by appealing to the distinction between personal and basic or generic interests that grounds the generic interest conception of the common good. Although individuals may adopt particular life plans that have little in common or that conflict or diverge in fundamental ways, each person who embraces such a life plan is committed to its value and, therefore, to the value of the freedom and capabilities necessary to formulate, pursue, and revise a first-order life plan. Despite differences in dress, demeanor, or aspiration, each person who takes the time to reflect can recognize every other person as a moral and political equal in this sense: to the extent that each is committed to a life plan grounded in some conception of the good, each is committed to the value of being able to formulate, pursue, and revise an individual life plan.

This shared higher-order interest in being free to advance one's personal interests defines the "space of equality," the domain over which all community members have a just claim to equal treatment. It is from this social standpoint that the institutions of research ethics are evaluated. As a result, the integrative approach is sufficiently general in scope that it applies to research carried out by individuals who fill a wide range of social roles, from physicians to economists, psychologists, management scientists, public health experts, citizen scientists, and anyone else seeking to generate socially valuable information from studies that involve human participants. Other frameworks that appeal to the narrow duties of particular professions lack this generality in scope.

The integrative approach also provides a unified framework for integrating various concerns that arise in this domain. The egalitarian research imperative is to foster a system of social cooperation in which individuals can take up, as part of their personal life plan, the project of developing the knowledge needed to better safeguard the basic interests of community members without sacrificing their status as free and equal persons. The justifications for exposing people to risk, and for limiting the risks that volunteers can accept, are grounded in the same shared interest of every individual in having real freedom to formulate, pursue, and revise a life plan of their own.

As we saw in §6.6, the integrative approach is capable of reconciling the social value requirement with the principles of equal concern, concern for welfare, rational expectation, and no impermissible gambles. Ensuring that trials reflect and address the uncertainty or disagreement in a moderately idealized community of experts ensures that research initiatives have a strong prima facie claim to producing socially valuable information while prohibiting study participants from being subjected to substandard care. This creates the conditions necessary for free and equal people to see research as an avenue through which they can generate the evidence that stakeholders need to more effectively, efficiently, and equitably meet the needs of community members without compromising their moral or social standing in the process.

As we saw in the previous chapter, other frameworks that seek to manage risk by using the template for the appeal to uncertainty rest on problematically narrow moral foundations. In particular, views grounded in the moral obligations of individual physicians are incapable of reconciling the social value requirement with the principles of concern for welfare or equal concern. Clinical equipoise fares better because it adopts a social conception of medical uncertainty. But it lacks generality because it cannot use the same normative foundation—the clinician's fiduciary duty to her individual patient—to provide guidance about how to evaluate all study risks. Rather, purely research-related risks have to be evaluated using a risk-benefit calculation of a kind that is inconsistent with the clinician's duty of individual care.

Component analysis recognizes this limitation and attempts to overcome it by presenting a comprehensive approach to risk assessment in medical research. But this expanded scope—the ability to cover interventions delivered with "therapeutic warrant" and those delivered solely for research-related purposes—is purchased at the price of conceptual disunity.

In component analysis, the risks from procedures that are offered with one motivation are bounded by the clinician's fiduciary duty while the only constraint on purely research-related risk is that it be outweighed by the value of the information a study is designed to produce. To the extent that the interests of participants is the normatively basic consideration, this appeal to the motivation or warrant for risk is morally arbitrary—whatever the source of a risk, consistency requires that the health and welfare of individuals be valued in the same way in relevant cases. If the welfare of participants is the morally basic concern, it seems arbitrary to circumscribe risks from interventions motivated by the prospect of direct benefit to the individual by the physician's duty of personal care and to allow the risks that are grounded in scientific necessity to be largely unbounded (see also Wendler and Miller 2007; Rid and Wendler 2010).

This inconsistent concern for participant interests is compounded by the fact that if an intervention is deployed with therapeutic warrant it does not follow that the condition being addressed is of sufficient medical importance that it affects the individual's basic or generic interests. Treatments that target a range of mild medical conditions (such as small cuts or abrasions, bruising, swelling, rash, temporary nausea, muscle aches, headaches, or male-pattern baldness) are delivered with therapeutic warrant. So, consistency would require that proponents of component analysis should hold that withholding a known effective treatment for such a condition in the context of a clinical trial would be unethical. In that case, component analysis faces two serious problems. First, it seems inconsistent to prevent participants from accepting risks from foregoing access to a known effective intervention for a minor medical condition, while permitting them to be exposed to significant risks from procedures that are necessary to collect study data. Second, this looks like an unacceptably paternalistic restriction since, for example, baldness seems to be a paradigm example of a medical condition whose meaning and significance will depend almost entirely on the contours of an individual's life plan.

Proponents of component analysis claim that they are not committed to such a position since non-treatment is a medically permissible response to minor medical conditions such as male-pattern baldness. But this appeal to professional practice is either an inappropriate deference to arbitrary professional authority or it is an unexplained explainer. It is an inappropriate deference to arbitrary professional authority if the normative ground for permitting the withholding of an effective intervention is the brute fact

that doctors find it acceptable not to treat this medical condition. In other words, if the normative basis of the appeal is to the preferences of medical professionals, as such, then it vests too much authority in those professionals since it empowers them to limit the decisions facing patients or study participants for reasons that are grounded solely in their preferences as professionals.

If, instead, the claim is that it is permissible for clinicians to leave baldness untreated because it is not a sufficiently significant health problem that it requires medical intervention, then we need an account of the severity or seriousness of medical conditions that is independent of the judgment of experts. But this is precisely what is lacking in views that appeal to the moral obligations of professionals for their normative foundation.

Differentiating standards of risk management on the basis of the warrant for deploying an intervention is also vulnerable to the charge of inconsistency from a different angle. In particular, the nature of the question that a study can answer is shaped by numerous variables including which interventions are provided to study participants out of the motive of therapeutic intent. If it is permissible to evaluate the risks of purely research-related interventions in terms of the value of the information the study is likely to generate, then it seems arbitrary to apply a different standard to other elements of a study that also impact the value of the information a study can produce. Since decisions about which diagnostic, therapeutic, or prophylactic interventions to provide to participants also affect the goal of producing scientifically sound and socially valuable information, critics argue that all aspects of a clinical trial should be assessed in terms of whether the associated risks are reasonable in light of the value of the information the study is likely to generate (Rid and Wendler 2010).

Proponents of component analysis have argued that there is a unified moral foundation underlying the different standards that are applied to these different components, namely, the importance of securing trust between study participants and the state (Miller and Weijer 2006b; Weijer, Miller, and Graham 2014). It is surely correct that a viable framework for assessing and managing research risks must be capable of securing social trust and I am deeply sympathetic to the claim that clinical equipoise should be thought of as an important mechanism for securing that trust. But within component analysis, this insight is swamped by the repeated claim that the central dilemma to be resolved is the reconciliation of the physician's duty to act in the patient's interest with the demands of clinical research and the assertion that

"clinical equipoise does not adequately specify the doctor-researcher's duty of care to the patient-subject" (Miller and Weijer 2006b, 546). As I argued in the last chapter, if the fundamental dilemma concerns reconciling research with the therapeutic obligations of clinicians, and if purely research-related interventions are delivered without therapeutic warrant, then importing a separate risk standard for such interventions reflects an unresolved internal tension within this approach.

In fact, the problems we have been rehearsing in this section are simply a reiteration of what I called the dilemma of determinate duties (§5.10.1). Views that appeal to the professional's duty of care require a standard for limiting risks that has sufficient content that it can provide useful guidance for regulating research risks. The traditional, Hippocratic interpretation of the duty of personal care provides such independent and operationally meaningful content but at the cost of being overly restrictive and unjustifiably paternalistic. Adopting a less paternalistic conception of the duty of care has the advantage of bringing the content of that duty into better alignment with the way it is understood within clinical medicine. The problem, however, is that it purchases this alignment at the cost of its independent, operational content.

In contrast, the integrative approach focuses directly on the interests of study participants. When individuals face risks to their basic interests, or experience conditions that restrict, impede, or impair those interests, those individuals are placed at a disadvantage. This is a disadvantage not merely with respect to goals or ends that they happen to have insofar as they, for example, value being a member of a particular club or aspire to look like a particular celebrity, but with respect to their ability to pursue their personal interests, whatever they are, on an equal footing with others. The integrative approach uses the condition of uncertainty regarding basic interests (§6.4.1) to ensure that study participants are not knowingly deprived of the means of securing their basic interests in the course of research. This means that asking participants to forego interventions that are intended to treat only minor ills is consistent with respect for those people as free and equal because those conditions do not affect their ability to formulate, pursue, or revise a reasonable life plan. In contrast, when patients face risks that threaten their standing in this space, it is inconsistent with respect for their standing as free and equal to provide them with a level of care that falls below what at least a reasonable minority of experts regard as best for their condition.

Like clinical equipoise and component analysis, the principle of uncertainty regarding basic interests (§6.4.1) invokes the judgment of experts. Unlike those views, the integrative approach only appeals to experts to determine whether an independent moral standard has been met. In this case, the standard is grounded in the special moral status of a person's basic interests. The judgment of experts reflects the state of medical knowledge about how best to safeguard or advance those interests.

Additionally, the integrative approach strives for consistency in the assessment and management of risks not just within research, but across other socially valuable activities. To the extent that other activities pose risks to the health and welfare of individuals, we should strive for consistency with respect to the treatment of like cases. As I argued in §6.3.3, a society that prohibits its members from taking risks to their basic interests would be oppressively restrictive in its limitation of legitimate life plans. Research is a social activity in which there are special considerations that warrant special forms of oversight, such as prospective review before committees of diverse membership (see chapter 7), but these should not preclude individuals from freely and knowingly accepting a broad range of risks to their personal interests and a reasonable range of affirmative risks to their basic interests.

6.7.3 The Belmont Approach

In the previous chapter I argued that the second dogma of research ethics is the view that research with human participants is an inherently utilitarian undertaking. In light of the profound problems with attempts to fill out the template for the appeal to uncertainty that rely on parochial moral foundations, some have argued that we should instead jettison that approach entirely and simply embrace the utilitarian essence of research.

This alternative has some distinct advantages. First, it has the advantage of founding risk assessments on a compelling normative foundation. In the *Belmont Report* (1979), the National Commission asserts that considerations of risk in research are grounded in the value of beneficence. Beneficence is attractive as a normative foundation for a framework of risk assessment since it reflects a fundamental concern for the welfare of persons.

Second, this compelling normative foundation is also sufficiently general in scope that it can be applied consistently to all aspects of research with humans. In particular, beneficence gives rise to a general duty that is

expressed by the rule "maximize possible benefits and minimize possible harms" (National Commission 1979, B.2). This concern for welfare thus requires consideration of both the risks and benefits that accrue to individual research participants as well as risks and benefits that accrue to society more broadly. As such, "beneficence thus requires that we protect against risk of harm to subjects and also that we be concerned about the loss of the substantial benefits that might be gained from research" (C.2).

By grounding questions of risk in the value of beneficence and embracing the language of maximization, the *Belmont Report* frames questions of risk as a kind of utilitarian optimization problem. Research risks are to be assessed and managed by quantifying the net impact of the potential benefits and risks to the welfare of individual study participants, quantifying the potential benefits to the welfare of the future persons who stand to benefit from the knowledge the study is designed to produce and then ensuring that the net gains to future persons are sufficient to outweigh any loss of welfare incurred by study participants. I will refer to frameworks that embrace this language of risk-benefit analysis as following the Belmont approach.

In the scholarly literature, the Belmont approach has been adopted and explicated within the non-exploitation approach (Miller and Brody 2002, 2003) and the net-risk approach (Wendler and Miller 2007; Rid and Wendler 2010), and it is one of the standards used in component analysis (Weijer 2000; Weijer and Miller 2004). As I argued in the previous chapter, proponents of the Belmont approach capitalize on the shortcomings of most frameworks that rely on a parochial foundation to fill in the content of the template for the appeal to uncertainty. Because those views appear to be the most natural and intuitive way to fill out the template and because they are riddled with deep problems, the Belmont approach appears to be the only coherent alternative.

At the end of the previous chapter, I argued that it is not necessary to think of research as an inherently utilitarian undertaking, and the main goal of the present chapter has been to establish how it is possible to reconcile the social value requirement with the principles of equal concern, concern for welfare, and no impermissible gambles. It is important to emphasize, therefore, that the utilitarian approach outlined in Belmont and refined by proponents of the net-risk and non-exploitation approaches is not somehow a necessary or privileged approach to risk assessment in this domain. Its connection to a long and well-developed moral tradition, and the appearance of technical and operational clarity adds to its allure. But I want to suggest that the

downsides of this approach are considerable and that its merits are not as substantial as they appear once you begin to consider them carefully.

First, the Belmont approach brings with it the specter of conflict that we discussed in chapter 2 between the rights and interests of study participants and the invariably greater good that stands to flow to future persons from advances in science and social progress. The single clearest aspect of the Belmont approach is the explicit moral permission that it provides to trade the welfare of study participants for sufficiently large increases in the welfare of future people. The most obscure aspect of this approach is whether this permission can be reconciled with a commitment to respect study participants as free and equal persons and whether respecting something like the principle of equal concern is even morally desirable within such a framework.

If the research enterprise is to be organized on terms that are capable of securing the voluntary cooperation of free and equal persons, then these shortcomings of the Belmont approach are not trivial. That framework foregrounds precisely the kind of moral permission that was used to justify past cases of abuse without having clear and coherent internal resources for reassuring community members that the institutions of research with human participants are designed to respect their status as free and equal. The Belmont approach thus lacks adequate resources for providing clear, public assurance that no stakeholder in the research enterprise will be subject to domination, abuse, exploitation, indifference, antipathy or wrongdoing at the hands of others.

To take one example, proponents of the non-exploitation approach argue that researchers are not bound by the clinician's duty of personal care but, instead, by the duty not to exploit study participants. In explicating the content of this requirement, they hold that it requires the observance of the following conditions:

(1) that risks to subjects are reasonable,
(2) that the research has social value and
(3) scientific validity,
(4) that subjects give free and informed consent,
(5) that there is fair subject selection,
(6) independent review, and
(7) respect for persons.

If these conditions seem familiar it is because they have been put forth as capturing the consensus requirements for ethically acceptable research with human subjects (Emanuel, Wendler, Grady 2000). As such several features of this proposal are striking.

First, the first three of these requirements are not actually constraints that limit the pursuit of utilitarian goals in clinical research. Rather, they serve to define those utilitarian goals and to express necessary conditions for their achievement. As such, these requirements would be endorsed by any legitimate utilitarian theory, whether it includes a constraint against exploitation or not.

Second, the other constraints are not limits on risk, per se. Because risk assessments have to be made prior to seeking informed consent, the requirement in (4) does not directly curb or limit the extent of the potential sacrifices that can be *asked* of trial participants in order to advance the common good. If we think of informed consent as the primary bulwark for limiting the risks to which study participants can be exposed in exchange for increases in the greater good, then inevitable defects in that process will result in research going forward that undermines the warrant for trust in the capacity of that system to respect study participants as free and equal persons.

Similarly, fair subject selection prohibits bias in the selection of who can be approached with the option to take on research risk, but that does not provide criteria for determining when those risks pass a limit or a threshold for what is reasonable. Independent review is an important safeguard, but keep in mind that the topic under consideration is what framework for risk assessment and management should govern that independent review process. So, we cannot appeal to independent review as an independent check on the framework for risk that is supposed to be used in that very review process.

Respect for persons, at least within the non-exploitation framework, covers a variety of protections for privacy, confidentiality, and the provision of information both during the conduct of research and once it is completed. This value does require careful monitoring of participant welfare, and "if subjects experience adverse reactions, untoward events, or changes in clinical status, they should be provided with appropriate treatment and, when necessary, removed from the study" (Emanuel, Wendler, and Grady 2000, 2707). But the ongoing monitoring of participant welfare is consistent with an ex ante study design that imposes significant burdens and exposes participants to significant risks for the purpose of advancing socially valuable and scientifically sound research.

The only principle for determining how individual health or welfare can be traded off against gains in knowledge, and therefore the advancement of the common good, is the requirement that risks be reasonable. Risks that are not offset by benefits to individual trial participants are judged to be reasonable if and only if they are sufficiently offset by gains in the knowledge that the research is designed to generate. At the end of the day, once we look carefully at the constituents of this view, the only substantive constraint on research risks is that they be justifiable in utilitarian terms.

My first concern is thus that even if the Belmont approach is a feasible alternative, by foregrounding the permission to trade participant welfare for benefits to future persons, it lacks clear resources for ensuring that the research enterprise is configured on terms that respect the status of all stakeholders as free and equal persons. Since the integrative approach grounds both the moral mission of research and the limits on the demands it can make of stakeholders in the same fundamental respect for persons as free and equal, it is preferable as a framework for regulating research risk.

My second concern with the Belmont approach is that it has a veneer of elegance and simplicity that obscures deeper problems that have yet to be addressed in research ethics. First, since this framework is grounded in concern for welfare, the risks and benefits in question should relate to the welfare of those affected. The value of the information a study is designed to generate must be represented as a function of the welfare of the various people likely to benefit from the information that will be generated. Similarly, the burdens and risks to study participants should be represented as a function of gains and losses to welfare incurred through study participation. Once these two quantities can be represented, whether their ratio is acceptable will depend on some normative standard or trade-off function. Each of these issues is left to deliberators to address at a largely intuitive level.

However, it is not clear that it is even possible to make the kind of interpersonal comparisons of welfare that this approach requires if it is to be taken seriously as a decision rule. Consider first that *intrapersonal* comparisons of welfare are facilitated by appeal to the way care advances or detracts from goals, ends, or means that are organized and ordered by an individual's personal life plan, given the ideals and values that motivate and suffuse it. When we consider whether the risks that an intervention poses to one person are outweighed by the expected benefits of that intervention for that same person, our ability to quantify and compare each side of this equation is facilitated by an understanding of the larger life plan and life projects of that

person. If the side effects of a medication make it more difficult for a patient to engage in activities that play a particularly central role in their individual welfare—in, for example, the subjective quality of their life or in their ability to function in ways that are central to defining projects or plans—then they may be willing to forego such care, or to undertake a less effective course of care without those specific side effects. As we saw at length in the previous chapter, in such patients, the optimal course of care, with respect to the patient's welfare, may be the course of care that is less effective or more burdensome when evaluated solely from the narrow standpoint of the patient's physical health. In contrast, other patients may have life plans in which their narrow health interests and the particular benefits and burdens of what is regarded as optimal clinical care from the standpoint of their narrow health interests dovetail with the contours of their larger life plans.

The key point is that each individual's larger life plan, and the conception of personal welfare that it helps to structure, creates a context in which different experiences and functionings can be compared and ordered because it is relative to that person's larger values, plans and projects that such experiences and functionings have determinate value. In other words, an individual's life plan creates the criteria or desiderata relative to which the benefits and burdens of various activities, including alternative courses of medical care, can be evaluated and ranked.

To make *interpersonal* comparisons of welfare, we have to compare the cumulative gains or losses to the welfare of one group of people to the cumulative gains or losses in welfare to another group of people (the population of study participants, on one side of the equation, and the larger beneficiaries of scientific progress, on the other side), all of whom may embrace different life plans. If we rely on the larger life plan of an individual to assign a determinate value to an experience or a functioning, then it is unclear how to make cross-life-plan comparisons of welfare. The life plan of each person may render welfare rankings or scores determinate and meaningful for that person, but cross-life-plan comparisons cannot be made relative to the contours and valuations of a single person's life plan.

Within economics, there is a history of skepticism about whether interpersonal comparisons of welfare are possible. Common sense as well as a long philosophical literature argues that we can make such comparisons because we often do make them. And these observations are sound, as far as they go. In particular, they are telling against a radical skepticism of the form that asserts that we cannot know anything about the welfare states of

others because we have no way to compare them to our own. But the question, for our present purposes, is not whether interpersonal comparisons of welfare are possible in any sense, but whether they are possible in the sense required by the calculous of risk we are considering here. And at this level, there appears to be room for significant doubt.

In particular, Louis Narens and Duncan Luce (1983) have shown that common sense is correct in holding that such comparisons are possible at least in the sense that pairs of individuals can, over time, develop a shared representation of each other's utility function. As a result, the common experience of being able to compare the magnitude of impacts on the welfare of those close to us to similar impacts on our own welfare can be vindicated at a technical level. What it appears we cannot do, however, is extend that shared representation to accommodate the utility functions of additional individuals. Since the calculous of risk required in research ethics must extend beyond pairs of individuals to groups, the skeptical position appears to hold at that level.

Although the results presented by Narens and Luce have far-reaching implications, they do not constitute a proof that it is impossible to make any kind of interpersonal comparison of welfare of the form required in research ethics. So, it might be argued that it is still reasonable to consider such comparisons to be possible. But this raises a second concern addressed by Kenneth Arrow (1951).

Arrow's concern is this: if it is possible to make such interpersonal comparisons of welfare, then it matters whether there is a single, unique way of doing this. If so, then the problem might be clarified over time and given ever greater clarity and precision. But if, as seems more likely, there are many different ways of making such comparisons, then we face a further decision about which of those ways we should use for the purposes of regulating research. In particular, because these different ways of representing the relative values in question can support making different trade-offs, our choice of welfare metric can implicitly influence the substantive decisions we make when we use that framework. So, it is not sufficient to establish that such comparisons are possible. We need to establish either that there is only one way to make them, or that of the many possible ways to make them, one approach is ethically correct or superior to the rest.

To the extent that all of this work is left to the intuitions of stakeholders, the concern is that the variety of different ways of specifying the value of scientific information and of trading it off against the interests of individual trial

participants are at least as vast as the imaginations of different deliberators. In other words, if it is possible to specify a metric for making such trade-offs, then there may be as many ways of doing this as there are different sets of weights that might be affixed to each kind of value.

Proponents of this or related approaches have dismissed these worries as a misguided desire for a kind of precision that cannot be had. Such issues ultimately boil down to matters of "judgment," they claim, and cannot be quantified (Miller and Brody 2002, 2003; Miller 2003). If we consider the role that risk assessments play in IRB deliberations, this is surely correct—there is no clearly worked-out utilitarian calculous in research ethics because stakeholders are not really making the computations that this kind of equation suggests. This is rather what Ruth Macklin has called a "pseudo-metric," a principle that is given a mathematical formulation but that is not part of any real formal system of assessment. In this case, the Belmont approach adopts morally troubling language that most clearly enunciates the permission of trading participant welfare for gains to the common good without any offsetting benefits that might come from the operational clarity of a precise formal system.

The integrative approach provides operationally meaningful guidance to stakeholders by setting clear criteria for making a prima facie claim to social value and then reconciling such research with a set of deeply compelling moral principles. Addressing uncertainty or conflict among experts about how best to advance the basic interests of community members helps to connect research questions with gaps in the ability of health systems to effectively, efficiently, or equitably meet important needs of community members. The importance of these needs is not cashed out in terms of welfare since welfare is a function of the substantive, first-order life plan of an individual. Instead, the importance of health needs is determined from the standpoint of an individual's shared, highest-order interest in being able to formulate, pursue, and revise a life plan of their own.

Whether smarter people than I can give more precise operational content to this distinction between personal and basic or generic interests remains to be seen. Until then, the integrative approach also rests on the informed and conscientious judgments of stakeholders. But, rather than enunciating the permissibility of sacrificing the interests of a few to promote the good of the many, this framework encourages stakeholders to design studies on terms that are consistent with respect for study participants as free and equal

persons. The judgments regarding risks and burdens required by this framework can be guided, in particular cases, by the underlying rationale for this distinction and by features of risk that are likely to track this distinction. These features include whether a harm is likely to be transient, temporary, or reversible or whether it is likely to be permanent and irreversible; whether it results in a limitation or impairment of ability to function or whether it affects the capacity of a person to perform a wide range of functions that are implicated in the process of forming, revising, and pursuing a reasonable life plan of one's own.

6.8 Conclusion

The integrative approach articulates the conditions under which it is possible to jointly satisfy the core requirements of the egalitarian research imperative. That imperative asserts a moral responsibility on the part of community members to create a system of research with humans that is responsive to the basic interests of community members in two ways. First, this division of social labor must be organized around the public goal of generating the evidence that is necessary to bridge the gaps between the basic interests of community members and the capacity of the basic social structures of that community to safeguard and advance those interests. Second, as a voluntary scheme of social cooperation, the research enterprise must be organized on terms that respect the status of all community members as free and equal persons.

The integrative approach also articulates the terms on which the production of socially valuable information can be reconciled with a network of compelling ethical requirements including the principles of equal concern, concern for welfare, and no impermissible gambles. This demonstrates that an imperative to carry out research can advance the common good without dispensing with the rule of law, without empowering anyone to make arbitrary judgments about the rights and welfare of others, and without running roughshod over the integrity of the individual.

If nothing else, the arguments of this and the previous chapter illustrate the importance of rejecting the problematic views that structure the conceptual ecosystem of orthodox research ethics, including the two dogmas of research ethics discussed in the previous chapter. Understanding research as a

social enterprise that serves a social purpose that is closely connected to the proper functioning of the basic social structures of a community (such as its individual or public health system) provides a solid foundation for ensuring that this cooperative enterprise is carried out on terms that its stakeholders can recognize as basically just.