Published in final edited form as:

Semin Fetal Neonatal Med. 2018 February ; 23(1): 68–74. doi:10.1016/j.siny.2017.10.003.

Neonatal research ethics after SUPPORT

John D. Lantos*

Children's Mercy Hospital, Kansas City, Missouri, USA

SUMMARY

The SUPPORT study (Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments), sponsored by the Eunice Kennedy Shriver National Institute of Child Health and Human Development to evaluate different oxygen saturation targets for extremely premature babies, led to a national controversy that was surprisingly public, intense, and polarizing. This article describes the study design, the study outcomes, and the key issues. I conclude that the controversy was based on two different views of the clinical investigator. One, held by investigators themselves, is that investigators are primarily committed to the patient's well-being. The other sees the investigator as unable to disentangle his conflicting loyalties and as inevitably prioritizing the goals of research over the goals of patient care. I suggest that our current oversight systems overstate the risks of research and understate the risks of idiosyncratic practice variation. A better system would treat the relative risks of these two phenomena as comparable.

Keywords

Neonatal; Research; Ethics; Regulation

1. Introduction

The controversy over the SUPPORT study (Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments) of oxygen saturation targets in extremely premature babies was surprisingly public, intense, and polarizing. In order to understand the controversy and its implications, I first describe the study design and the study outcomes, then review the issues that led to controversy, and finally discuss the implications.

2. SUPPORT study design

Between 2005 and 2009, infants born between 24 and 27 weeks of gestation were randomized to two different oxygen saturation targets. One group was randomly assigned to

* Address: Children's Mercy Hospital, 2401 Gillham Road, Kansas City, MO 64108, USA. Tel.: +1 816-701-5284. jlantos@cmh.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest statement

The author was retained as an expert witness by a law firm that was defending Masimo Corporation, the company that manufactured the oximeters used in the SUPPORT study.

a target range of oxygen saturation of 85–89% (the lower-oxygen-saturation group). The other group was randomized to 91–95% (the higher-oxygen-saturation group).

This targeting of oxygen saturation was initiated within the first 2 h after birth and was continued until 36 weeks of postmenstrual age or until the infant was breathing ambient air and did not require ventilator support or continuous positive airways pressure for >72 h. The target ranges were kept unchanged from birth until 36 weeks of postmenstrual age.

Blinding was maintained with the use of electronically altered pulse oximeters. For babies in the low-oxygen group, an oximeter reading of 90% corresponded to actual levels of oxygen saturation of 87%. For babies in the high-oxygen group, that same reading corresponded to an actual saturation of 93%. The oxygen-saturation reading gradually changed and reverted to actual values when it was <84% or >96% in both treatment groups. Doctors, nurses, and respiratory therapists were then asked to maintain oxygen saturations between 88% and 92% in all babies, knowing that, by doing so, some would have oxygen saturations of 85–89% and others would have saturations of 91–95%.

The primary outcome of the study was a combined variable of severe retinopathy or death. All surviving infants were followed by ophthalmologists. Eyes were examined beginning at 33 weeks of postmenstrual age and continued until the study outcome was reached or resolution occurred. Retinopathy was treated using standard protocols which could have included laser therapy, cryotherapy, both laser therapy and cryotherapy, scleral buckling, or vitrectomy.

An independent data and safety monitoring committee reviewed results at approximately 25%, 50%, and 75% of planned enrollment.

The study was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development. It was approved by the institutional review board at each participating site. Written informed consent was obtained from the parent or guardian of each child before delivery.

3. Study results

The study results were published in two papers in the *New England Journal of Medicine* (NEJM) [1]. The first paper, in 2010, reported short-term outcomes. There was no difference in the rate of the combined outcome of “severe retinopathy or death before discharge” between the lower- and higher-oxygen-saturation groups (28.3% and 32.1%, respectively). There was, however, a statistically significant difference in death before discharge. This occurred in 130/654 (19.9%) of infants in the lower-oxygen-saturation group compared with 107/662 (16.2%) of infants in the higher-oxygen-saturation group (relative risk (RR) with lower oxygen saturation: 1.27; 95% confidence interval (CI): 1.01–1.60; $P = 0.04$).

The rate of severe retinopathy among survivors was lower in the lower-oxygen-saturation group (8.6% vs 17.9%; RR: 0.52; 95% CI: 0.37–0.73; $P < 0.001$). Babies in the lower-oxygen-saturation group were less likely to require supplemental oxygen at 36 weeks post-conceptual age.

In a follow-up paper in 2012 [2], researchers reported that there were no statistically significant differences in the rates of neurodevelopmental impairment at 18–22 months between the groups (30.2% vs 27.5%; RR: 1.12; 95% CI: 0.94–1.32; $P = 0.21$).

4. The controversy

The controversy began on March 7th, 2013, when the federal Office for Human Research Protections (OHRP) notified University of Alabama at Birmingham (UAB) that it had concerns about the informed consent forms in the study [3]. They found fault with the consent form for not including information about “prior research and analyses that had been done looking at the relationship between oxygen and ROP [retinopathy of prematurity]” and not identifying “any specific risk relating to randomizing infants to a high or low range of oxygen.” This, OHRP claimed, was in violation of federal regulations that require all informed consent forms for research to include “a description of any reasonably foreseeable risks and discomforts.” OHRP requested that UAB provide OHRP with a plan to improve the ways in which they would “adequately address the basic elements of consent as required by HHS [Health and Human Services] regulations.”

A month later, a Washington-based lobbying group, Public Citizen, wrote to the Secretary of Health and Human Services alleging that OHRP had not gone far enough [4]. Public Citizen claimed that “any study comparing the two experimental target levels of oxygen saturation would be both unethical and not compliant with requirements of HHS regulations.” They demanded that HHS issue “a formal apology to the parents of all 1316 infants who participated in the SUPPORT study.”

A few days later, on April 15th, an editorial in the *New York Times* asserted that the SUPPORT study “failed to meet the most basic standard: providing an informed consent document to parents that accurately described the risks and benefits of the research to be conducted.” [5]

Two days later, an editorial in the *NEJM* defended the study, the consent form, and the investigators. Editor-in-Chief Drazen wrote that the consent document for the SUPPORT study “... addressed the prevalent knowledge fairly and reasonably.” He claimed that, at the time that the study was designed, “there was no evidence to suggest an increased risk of death with oxygen levels in the lower end of a range viewed by experts as acceptable, and thus there was not a failure on the part of investigators to obtain appropriately informed consent from parents of participating infants.” Drazen concluded that the investigators were being faulted with not knowing (and therefore disclosing) a risk that was not known until the study showed it to be present. “This is how new medical knowledge is gained,” he concluded.

In the same issue of the *NEJM*, two prominent bioethicists defended the study and argued that OHRP’s actions were wrong. They wrote, “The OHRP is asking that research be described as riskier than it really is and is suggesting that the parents were duped into enrolling their frail infants in dangerous research. Not only is that not true, but it also poses substantial risk to the conduct of valuable comparative effectiveness research both for

premature infants and for the general public who continue to face too many treatments where uncertainty prevails about what is best.” [6]

The bioethics community was divided in their response to the controversy. One group of ethicists argued that, “The infants included in the study were randomly assigned to oxygen-saturation targets that were consistent with standard clinical care at the participating institutions. The conclusion of the OHRP that the study’s experimental evaluation of these otherwise routinely used oxygen-saturation levels exposed subjects to additional risk (above the risks of routine clinical treatment) is not supported by the evidence.” [7] Another group came to the opposite conclusion [8]. They wrote, “The potential risks and benefits of being in the study could not be said to be the same as the potential risks and benefits of receiving care outside the study.” They concluded that the consent forms were deficient and thus that OHRP’s actions were appropriate. Neither group addressed Public Citizen’s more wide-ranging critiques of the study itself.

Leaders of the National Institutes of Health (NIH) defended the study and the consent process. They wrote, “The investigators (in SUPPORT) had no reason to foresee that infants in one study group would have a higher risk of death than would those in the other group. The babies included in SUPPORT were, of course, facing substantial risks because of prematurity – the same risks as premature babies who were not enrolled in the study – but their care was never compromised for the sake of the study.” [9]

At that point, OHRP withdrew its finding that the consent process for SUPPORT was flawed. They acknowledged “widespread misunderstanding of the risks that are required to be disclosed in obtaining informed consent for certain types of clinical trials.” [10] In order to help resolve those misunderstandings, OHRP called for an open meeting at the Department of Health and Human Services. At that meeting, 28 speakers offered comments to a panel of leaders from NIH, OHRP, and Food and Drug Administration [11,12].

A year after this meeting, OHRP issued a new “Draft Guidance on disclosing reasonably foreseeable risks in research evaluating standards of care” and asked for public comments [13]. The public was divided [14]. To date, OHRP has not issued a final version of those guidelines, though some of the concerns were addressed in recently issued revisions to the Common Rule [15].

4.1. Legal controversy

In the meantime, a lawsuit was filed on behalf of three children who had been enrolled in the SUPPORT study. The plaintiffs alleged that, “As a result of the careless, negligent, and reckless conduct of the defendants, the plaintiffs and the class were caused to suffer excruciating and agonizing pain, physical discomfort and emotional distress.” [16]

The lawsuit was eventually dismissed by a summary judgment [17]. The standards for granting summary judgment in a case like this are very high. The judge, Karon Bowdrie, was required by law to view all of the facts “in the light most favorable to (the plaintiff),” and then decide whether “... reasonable minds could differ on the inferences arising from undisputed facts.” [18] If so, then the court would deny summary judgment and let the case

go to a jury. In deciding that reasonable minds could not differ, then, the judge concluded that there was essentially no compelling evidence that plaintiffs were harmed by their participation in the study.

That is the history. The rest of this article examines the historical background and proposes some lessons that we could learn from this controversy.

4.2. Fundamental issues

One fundamental issue underlying the initial controversy could almost be classified as a technical issue of epistemology. It turned on a phrase in the Common Rule about the proper way to obtain informed consent for research. Investigators are required to provide “a description of any reasonably foreseeable risks or discomforts to the subject.” [19]

So, in one sense, the controversy was about whether researchers failed to disclose “reasonably foreseeable” consequences of participating in the study. That would have been bad. And the question of whether there were reasonably foreseeable harms that should have been disclosed and were disclosed not is important.

But that issue is not as important as an issue that arose later in the controversy when an organization called “Public Citizen,” a Washington DC-based lobbying group, upped the ante by claiming that, contrary to OHRP’s stated opinions, the problems of the SUPPORT study went far beyond the technicalities of consent. Instead, Public Citizen claimed, the study was unethical in its design. Public Citizen claimed that any study that involved randomization of babies to two different oxygen saturation targets was unethical and should never have been approved. By this view, it would have been both illegal and unethical to have even sought consent for such a study. Public Citizen called upon the Secretary of Health and Human Services to issue a formal apology to the parents of every baby who had been enrolled in the study. This call for a public apology was an implicit reference to President Clinton’s apology to the men who participated in the Tuskegee Syphilis study, which is widely viewed as one of the most egregious violations of research ethics in American history.

The two claims – that of OHRP and that of Public Citizen – share a common theme. Both turn on questions about what doctors knew, or should have known, or could have known, at the time they designed the SUPPORT study. If they knew that being in the study would expose babies to risks that they would not have been exposed to outside of the study, then those risks should clearly have been disclosed. But Public Citizen’s logic is exactly right. If they knew that participation in the study carried risks that could be avoided by not participating, and did not have any chance of offering benefits, then the study should never have been approved or carried out in the first place. If, on the other hand, doctors truly did not know what oxygen saturation targets were associated with the best outcomes for babies, then it would have been appropriate to do the study and to claim, as the consent form claimed, that there were no reasonably foreseeable risks associated with being in the study.

In order to evaluate the various claims and to understand the implications of this controversy, I will break down the controversy into seven separable concerns.

4.3. Six concerns about clinical research

Six separable aspects of prospective studies of oxygen make these studies controversial. These are: attributable risk; randomization, treatment by protocol; stopping rules; lack of a “standard care control arm”; and the need for, and techniques for, masking of investigators regarding treatment assignment.

4.3.1. Attributable risk—Premature babies are known to have a predictable likelihood of developing certain medical problems – including death, neurodevelopmental disability, and retinopathy. Each of these problems can occur with or without the other problems. Risk, in this context, means the population- or group-level prevalence of developing these problems.

In such a setting, in which conventional treatment is predictably associated with serious risks of mortality and known morbidities, it is controversial to decide which aspects of a study are likely to increase that underlying risk, decrease it, or have no effect on it. This assessment of “attributable risk,” that is, the risk that ought to be attributed to study participation as opposed to being attributed to the clinical condition of the baby and the known risk of conventional treatments, is important because any meaningful informed consent for research must disclose to parents the risks of study participation.

Can one assume that it is riskier to be in a study in which babies are randomized to one treatment or another when both of the treatments are in widespread use and a baby’s chance of getting one or the other outside of the study is comparable to their chance within the study? Most statisticians would say no. Many regulators and bioethicists would say yes. The answer turns, it seems, on the confidence that one has in the individualized clinical judgment of the doctors who are treating the patient. If they know that one treatment is better than a protocol that offered half the babies a treatment that was known to be worse, that would, of course, add risk. But if we assume that they do not know which treatment is better, as we must if we consider the study to be well-designed, ethical, and worth doing, then the attributable risk of the study is minimal or nonexistent.

4.3.2. Randomization—Randomization has been ethically controversial since it was first proposed as an important element of study design in the 1930s. Two decades ago, Hellman and Hellman wrote that randomized trials raise “the classic conflict between rights-based moral theories and utilitarian ones.” [20] Further, they suggest that this conflict is inherent and irreducible because it is part of the fundamental nature of clinical trials. By this view, patients have rights that are safeguarded by their uncompromisingly loyal clinician. When, however, they enroll in a research trial, their physician becomes a clinical investigator and thus, in essence, a double-agent. The clinician-investigator must balance to obligation of loyalty to the individual patient with an obligation to do good research. The Hellmans write, “Researchers are required to modify their ethical commitments to individual patients and do serious damage to the concept of the physician as a practicing, empathetic professional who is primarily concerned with each patient as an individual.”

Ethical concerns about randomization turn on one fundamental question – does the doctor know what is best for the patient or is the doctor truly uncertain? If one treatment is known to be superior, then randomization to a treatment that was known to be inferior would be

unethical. A doctor should refuse to participate in a clinical trial by recruiting his patients into the trial under those circumstances. If, however, both the individual doctor and the leading medical experts are truly uncertain about which treatment is better, then randomization may be ethically acceptable. In these circumstances, the act of assigning patients to different treatments randomly would not compromise the physician's fiduciary loyalties to the patient.

That all sounds good in theory. It is more difficult in practice because doctors often think that they know which treatment is better, even if the evidence to buttress their clinical judgment is lacking. Nevertheless, doctors have to treat patients and, in doing so, they often develop protocols based on the best available evidence, even if the quality of that evidence is not great. If, furthermore, the professional uncertainty about the best treatment is so widespread that expert physicians have come to different conclusions and accordingly had decided to use different protocols in their day-to-day practice, then the risk of randomization would be comparable to the risk inherent in the variation of actual clinical practices that flowed from the disagreements within the professional community. Whereas random treatment assignment will shift the treatments of some patients within the range of currently accepted standards of care, due to the uncertainty regarding which treatment is superior (which is the essential presupposition of clinical trials) no patient or physician would be sure whether this random shift in treatment would lead to benefit or harm for either group of patients and thus for any individual patient within those groups.

4.3.3. Treatment by protocol—The issues of treatment by protocol are related to the issues of randomization. In the SUPPORT study, the intervention required a series of many adjustments of treatment over days or weeks or months.

To adhere to the protocol, doctors could not make adjustments in oxygen based on their clinical judgments as to how much oxygen the babies needed. Thus, they gave up their ability to make individualized assessments of what was best for each baby.

In order to know whether this represents a problematic ethical compromise, it is necessary to know what the standard of care for oxygen therapy was at the time. If doctors routinely made individualized assessments of oxygen and fine-tuned the targeted saturation based on the baby's clinical situation, then the research protocol would represent a troubling deviation from standard care. If that were the case, one would question whether the study was ethical at the outset, since it would seem to be a study in which all babies were treated in a manner that fell short of prevailing standards of care.

Better understanding of the situation in neonatology at the time reveals that, here, too, babies are unique. Oxygen therapy in the neonatal intensive care unit (NICU) differs in important ways from standard oxygen therapy in other clinical circumstances. Because oxygen is uniquely toxic for babies, NICU doctors cannot rely on their individualized clinical judgment to decide what level of oxygen is best. Such an approach would always lead them to give more rather than less oxygen. And more, rather than less, is not always best. Thus, neonatologists, as part of their routine care, always use protocols. They decide, based on the

best available evidence, which is the appropriate saturation target, and then they try to keep the baby's own level as close to that target as they can.

Neonatologist Keith Barrington explained this phenomenon as follows:

Some centers felt that a saturation of 90 to 95% was the best idea, largely avoiding hyperoxia, and not wanting to risk hypoxia. So they would have a unit routine. The alarm limits for all preterm babies in the NICU would be set to, say, 89 and 96. Then the nurse would adjust the oxygen delivered, sometimes every couple of minutes, to stay within the target range. ... Every preterm baby in the NICU would have the same limits, until they no longer needed oxygen. Another center examining the same data would set their oximeter alarm limits to a lower range. Within each NICU, doctors would agree, based on their review of the evidence, about the target saturations that they would try to achieve. [21]

4.3.4. Stopping rules—The SUPPORT study had a data safety monitoring board (DSMB). The DSMB periodically compared the rates of serious adverse events between the two arms of the study to see if they were statistically significant.

In the SUPPORT study, the mortality rates at discharge from the hospital were 130 out of 654 (19.9%) in one arm and 107 out of 662 (16.2%) in the other. This just crossed the threshold of statistical significance (95% CI: 1.01–1.60; $P = 0.045$). The mortality rates at 36 weeks of post-conceptual age, by contrast, were 114 out of 654 (17.7%) and 94 out of 662 (14.2%). This was not statistically significant. A DSMB might have looked at both and would then have made a judgment call about whether to continue or stop the study.

When the SUPPORT study was published, the DSMB for the BOOST trial in Australia, New Zealand and the UK examined their interim data. They reported, in a letter to the *NEJM*, that their interim analysis did not reveal a statistically significant difference in mortality. They thus elected to continue the study. They wrote:

The independent data and safety monitoring committee in the Australian Benefits of Oxygen Saturation Targeting II (BOOST-II) trial and the New Zealand BOOST-NZ trial, whose protocols are similar to those in SUPPORT, has reviewed outcomes in 1352 patients. The committee reported no clear difference between the two oxygen targets in terms of the rate of death before hospital discharge. When considering the Australian BOOST-II and New Zealand BOOST-NZ trials in combination with SUPPORT, the committee found that there was significant uncertainty about the effects of treatment on mortality. They recommend that recruitment of study subjects continue. To our knowledge no such randomized controlled trial has reported survival free of disability in childhood, the major end point of all current oxygen-targeting trials. Until these results are known it would be premature to adopt either the higher or lower oxygen target for use in routine care. [22]

These analyses and decisions show that the investigators and their data monitoring boards were analyzing mortality data independently, even though the difference in mortality was not one of the primary endpoints of the study.

There is no completely objective way to determine whether a study ought or ought not to continue. If there were, then there would be no need for DSMBs. DSMBs must make difficult judgment calls when interim analyses show worrisome trends that have not yet – or have just barely – reached statistical significance. That was done in this study.

4.3.5. Lack of a concurrent “standard care” control arm—Some critics of SUPPORT focus on the lack of a “standard care” control arm. They claim that, because the standard of care at the time was to target an oxygen saturation of 85–95%, that the randomization to the low end and the high end of that range was ethically problematic. Dr Charles Nathanson argued that, “randomizing the critically ill to extremes of titrated therapies, creates practice misalignments which carry risks and do not represent usual care.” [23]

Among experts in research methodology, there is a vigorous debate about whether a standard care control arm is essential in order to evaluate the efficacy of variations in standard care. To a certain extent, the debate, as it applies to SUPPORT, turns on questions of what standard care actually was in NICUs around the world for babies born at 24–27 weeks gestation. Nathanson assumes that, because the recommended treatment was to target an oxygen saturation of 85–95%, this was both the agreed-upon standard and the most usual practice. This assumption is likely wrong. Many textbooks of neonatology recommended other target saturations. Many published papers recommended other targets. Many neonatologists chose other targets as their standard NICU protocols.

Rather than imagining that the conventional treatment at the time was to target 85–95%, a more accurate statement would be that, in each NICU, there was a different target and most of those targets were within the range of 85–95%. Some, as part of their standard practice, targeted the lower end of the range. Others targeted the higher end. To study each of these treatments and compare one against the other would have been methodologically impractical. At the very least, such a study would have been much larger, requiring not just two arms, or the addition of an elusive and illusory standard of care arm, but with many more arms to reflect the widespread and manifold variation in standard treatment.

Meade and Lamontagne point out three concerns with usual-care control arms. First, there is the question of how broadly to define “usual care.” Some centers in multicenter studies may not meet established standards of care. Thus, with SUPPORT, the standards in 2005 were to treat toward a target oxygen saturation of 85–95%. Some centers allowed lower oxygen saturations. Others targeted higher saturations. Should such centers have been included in the usual care arm? If so, should they have been instructed to change their “usual care” practices and to treat within the target range of 85–95%? If so, that would have been a deviation from “usual care,” necessitating yet a fourth arm. If usual care was not standardized in some way, the study results would likely have been uninterpretable. The study certainly would have taken much longer to complete, as it would have required many more subjects.

Second, as noted above, “usual care” in neonatology generally was provided by a predetermined protocol. Many critics of SUPPORT imagine that “usual care” would be to

provide individualized adjustments of oxygen levels based upon the baby's clinical condition from moment to moment. But that is not, as we have shown, usual care in neonatology.

A third problem with "usual care control arms" is that usual care changes over time. The longer a study takes to complete, the more "usual care" might change from the beginning to the end of the study. We know, for example, that practices with regard to oxygen therapy were in flux during the SUPPORT years. The longer the study went on, the more changes would likely have occurred in usual care. To the extent that usual care changed to reflect the experiences of clinicians who were simultaneously involved in a prospective study, it likely would have come to resemble the study treatments and have been more difficult to distinguish from them.

A final argument against the need for usual care control arms is that there are often good historical data as well as concurrent (but not randomized) data from similar patients who are not enrolled in the study. With regard to premature babies, there were excellent historical data from well-established national databases that allowed comparison of study babies with babies who were treated outside of the study. Of course, there are known problems with non-randomized comparison cohorts. But those problems do not disappear when study subjects are randomized. Instead, each comparison group gives valuable information.

4.3.6. Masking by use of altered oximeters—Many critics were concerned about the altered oximeters that were designed to give slightly false readings of oxygen saturations. Masking investigators and patients is a standard technique for randomized controlled trials.

Most of the critiques of the altered oximeters focused on the ways in which they prevented doctors from exercising individualized clinical judgment. As we have shown above, however, this concern is misplaced in the context of the NICU where oxygen supplementation is seldom administered based upon individualized clinical judgment. The masked oximeters did not prevent doctors from treating patients as they would have done anyway, that is, by protocol. Instead, the masked oximeters simply allowed them to do so with the sort of masking that is well accepted in randomized trials.

4.4. Lessons learned

By this analysis, it would appear that the SUPPORT study was well designed, carefully conducted, attentively monitored, and responsibly published. There is no evidence that babies were harmed by participating in the study. The results of the study supply doctors with important information that will help babies in the future. It seems to have been a model of ethical research on a difficult clinical conundrum in a particularly vulnerable population.

Given those conclusions, the researchers should have been honored and congratulated for their work to make the medical care of babies safer and more effective. Instead, they were cited for ethical violations by the federal government, pilloried by advocacy groups and the mainstream media, and hauled into court to defend themselves against perceived violations of parents' rights. Why?

The perceptions of the study reflect a deep-seated distrust of research and researchers that is central to research regulation today. Churchill was, perhaps, the first to articulate the concerns that undergird this distrust. He saw researchers as torn by dual loyalties. He wrote, "The acknowledged goal of the physician–patient relationship is healing or the health of the patient. The scientific investigator cannot claim this goal or the moral authority which goes with it." He saw these goals, and the relationships that developed from them, as mutually exclusive. "The two relationships have lives of their own which, by their very nature, compel or urge to certain priorities and inclinations to perceive and act in certain ways." [24] OHRP echoed this line of thinking. They wrote, "Ultimately, the issues in this case come down to a fundamental difference between the obligations of clinicians and those of researchers. Doctors are required, even in the face of uncertainty, to do what they view as being best for their individual patients. Researchers do not have that same obligation." [25] Bioethicists Macklin and Shepherd agreed, "It is the doctors, not the researchers, who have a fiduciary obligation and long-standing ethic to pursue the patient's best interests above all other considerations." [26] Annas echoed this view, "A physician must be guided by a fiduciary obligation to the patient. A researcher has no such obligation." [27]

Joffe and Miller took these arguments to a logical conclusion. They suggest not only that the motivations of researchers lead them to choices that compromise the interests of their study subjects. They suggest that the very nature of research morally obligates the researcher to do so. They write:

In the context of medical care ... beneficence entails the obligation of the health care provider to do what is best medically for particular patients, and nonmaleficence entails the obligation to avoid exposing patients to harms or risks of harm that are not compensated by the prospect of direct medical benefits. In clinical research, in contrast, beneficence directs investigators to promote social value by generating scientific knowledge, in addition to promoting medical benefit for subjects when consistent with the design of the research; nonmaleficence requires that risks are minimized and justified by the potential value of the knowledge to be gained by the research. [28]

By these views, the clinician is a virtuous deontologist, guided by firm moral principles to always, unfailingly, guard the interests of her patients. The researcher, by contrast, is an idealistic utilitarian, guided by equally firm moral principles to prioritize the interests of good science over the interests of current patient-subjects.

4.5. An alternative view: researchers as patient advocates

Most researchers do not see themselves in this way. Instead, they see well-designed research as entirely harmonious with uncompromised loyalty to the best interests of their patients.

Barrington articulates this view. He notes, "I have a fiduciary obligation to provide optimal treatment. I also have a moral obligation to know what the optimal treatment is. I also, simultaneously, have a moral obligation as a researcher to keep trying to find out what the best treatments may be." [29]

This view, too, has deep roots. Jay Katz, a pioneer of research ethics, knew that research and therapy overlap in complex ways, “The multiple purposes of medical practice, caring for patients, advancing science, improving the health of the community, nations, and future generations cannot be separated clearly.” He concluded that “research and therapy, pursuit of knowledge and treatment, are not separate but intertwined.” [30] Passamani sees these overlapping roles as a way to “protect physicians and their patients from therapies that are ineffective or toxic.” Grunberg and Cefalu posit that the alternative to such clinical research is not individualized and thus better patient care, but is merely a pretense of omniscience that physicians do not and cannot possibly have [31].

Similarly, Kass and colleagues argued that well-designed research studies are often safer than clinical treatment outside of studies. They wrote, “There is no good evidence to support the empirical assumption that research studies, as a class, are more likely than clinical practice to run counter to the medical best interests of patients.” [32]

4.6. Finding common ground

The debate between these two views of the clinical investigator has implications for the way we regulate clinical research. If the clinical investigator is viewed as reflective, trustworthy, and capable of maintaining his primary commitment to the patient’s well-being, even when the patient is enrolled in a research study, then assessment of the risks of research will – and should – focus primarily on the risks of the study procedures themselves. If, on the other hand, the investigator is seen as unable to disentangle his conflicting loyalties and as inevitably prioritizing the goals of research over the goals of patient care, then careful and constant oversight will be necessary even when the research procedures themselves entail no risk over and above the risks of treatment.

Our current system of research regulation reflects the latter view. It treats researchers as incapable of making even simple moral judgments about study design, patient enrollment, informed consent, or even data analysis.

This approach fails to protect patients because its focus is wrong. It leads to a world in which patients are exposed to the risks of inadequately studied innovative therapy, unjustified practice variation, and the divided loyalties of their clinicians. It mandates that researchers who are engaged in activities that cause no risk to patients be hampered by burdensome regulations. The fundamental change that is required in this system is not one that tinkers around the edges [33]. It requires a fundamental theoretical and conceptual shift. We must acknowledge that research can be done primarily to benefit current patients and that researchers are generally able to balance their moral obligations to those patients.

This does not, of course, obviate the need for informed consent. Instead, it would demand a consent process that more accurately conveys the risks and benefits of both being in a clinical trial and of not being in such a trial. It would honestly portray the fact that we do not know which babies in which study arm will do better or worse, and that we do not know whether babies in the study will do better or worse than babies who are not in the study.

New federal regulations recognize this need. Recently released revisions to The Common Rule note that “consent forms have been growing longer and can be difficult to understand. They too often appear to be designed more for protecting the legal interests of institutions conducting research than for helping someone make a decision about participation.” [34] They recommend that this be changed and that, “Prospective participants [must] be given the information that a “reasonable person” would want to have in order to make a decision about participating.”

This is an important change. Short and simple consent forms are more likely to empower patients than long forms that include a lot of information that is not directly relevant to the choice that parents have to make.

Acknowledgments

Funding sources None.

References

1. SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Carlo WA, Finer NN, Walsh MC, et al. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med*. 2010; 362:1959–69.
2. Vaucher YE, Peralta-Carcelen M, Finer NN, et al. Neurodevelopmental outcomes in the early CPAP and pulse oximetry trial. *N Engl J Med*. 2012; 367:2495–504. [PubMed: 23268664]
3. Letter from Office for Human Research Protections to University of Alabama at Birmingham. [http://www.hhs.gov/ohrp/sites/default/files/ohrp/detrm_ltr/YR13/mar13a.pdf]
4. Public Citizen letter to Secretary of HHS. <http://www.citizen.org/documents/2111.pdf>
5. Editorial Board of the New York Times. An ethical breakdown. <http://www.nytimes.com/2013/04/16/opinion/an-ethical-breakdown-in-medical-research.html>
6. Magnus D, Caplan AL. Risk, consent, and SUPPORT. *N Engl J Med*. 2013; 368:1864–5. [PubMed: 23597408]
7. Wilfond B, Magnus D, Antommaria AH, et al. The OHRP and SUPPORT. *N Engl J Med*. 2013; 368:e36.
8. Macklin R, Shepherd L, Dreger A, et al. The OHRP and SUPPORT – another view. *N Engl J Med*. 2013; 369:e3. [PubMed: 23803134]
9. Hudson KL, Guttmacher AE, Collins FS. Support of SUPPORT – a view from the NIH.
10. Office for Human Research Protections letter to University of Alabama at Birmingham. Jun 4. 2013 https://www.hhs.gov/ohrp/sites/default/files/ohrp/detrm_ltr/YR13/jun13a.pdf
11. Transcript of Meeting at HHS, August 28th, 2013 [<http://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/public-meeting-08-28-2013-transcript/index.html>].
12. https://www.youtube.com/results?search_query=hhs+public+meeting+august+2013
13. Office for Human Research Protections draft guidance. [<http://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/draft-guidance-disclosing-risk-in-standards-of-care/index.html>]
14. Comments on Office for Human Research Protections draft guidance. <https://www.regulations.gov/docketBrowser?rpp=25&so=DESC&sb=commentDueDate&po=0&dct=PS&D=HHS-OPHS-2014-0005>
15. Menikoff J, Kaneshiro J, Pritchard I. The common rule, updated. *N Engl J Med*. 2017; 376:613–15. [PubMed: 28103146]
16. Oliver, M. Lawsuit says UAB study harmed premature babies. http://blog.al.com/spotnews/2013/04/lawsuit_says_uab_study_harmed.html
17. Lantos J. Vindication for SUPPORT. *N Engl J Med*. 2015; 373:1393–5. [PubMed: 26332442]

18. Moore, Looney v. Case No.: 2:13-cv-00733-KOB United States District Court for Northern District of Alabama. Aug 13. 2015 http://www.gpo.gov/fdsys/pkg/USCOURTS-alnd-2_13-cv-00733/pdf/USCOURTS-alnd-2_13-cv-00733-0.pdf
19. Code of Federal Regulations. Protection of Human Subjects 45 CFR 46. <http://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/>
20. Hellman S, Hellman DS. Of mice but not men – problems of the randomized clinical trial. NEJM. 1991; 324:1855–9.
21. Barrington KJ. Personalized medicine in the NICU. Am J Bioeth. 2013; 13:33–5.
22. Tarnow-Mordi WO, Darlow B, Doyle L. Target ranges of oxygen saturation in extremely preterm infants. N Engl J Med. 2010; 23:363–1285.
23. Nathanson, C. Testimony at HHS Open Meeting. Aug 24. 2013 <http://www.hhs.gov/ohrp/newsroom/rfc/Public%20Meeting%20August%2028,%202013/supportmeetingtranscriptfinal.html>
24. Churchill L. Physician-investigator/patient-subject: exploring the logic and the tension. J Med Philos. 1980; 5:215–24. [PubMed: 7205087]
25. Letter from Office for Human Research Protections to University of Alabama at Birmingham. http://www.hhs.gov/ohrp/detrm_lettrs/YR13/jun13a.pdf
26. Macklin R, Shepherd L. Informed consent and standard of care: what must be disclosed? Am J Bioeth. 2013; 13:9–13. [PubMed: 24256523]
27. Annas, G. Presentation for HHS Public Meeting on Protection of Human Subjects. <http://www.regulations.gov/#!documentDetail:D=HHS-OPHS-2013-0004-0034>
28. Joffe S, Miller FG. Mapping the moral terrain of clinical research. Hast Center Rep. 2008; 38:30–42.
29. Barrington, K. What are the responsibilities of clinical researchers?. <http://neonatalresearch.org/2013/09/18/what-are-the-responsibilities-of-clinical-researchers/>
30. Katz J. The education of the physician-investigator. Daedalus. 1969; 98:480–501.
31. Grunberg SM, Cefalu WT. The integral role of clinical research in clinical care. N Engl J Med. 2003; 348:1386–8. [PubMed: 12672868]
32. Kass NE, Faden RR, Goodman SN, et al. The research–treatment distinction: a problematic approach for determining which activities should have ethical oversight. Hastings Cent Rep. 2013; (Suppl 1):S4–15. [PubMed: 23315895]
33. Emanuel EJ, Wood A, Fleischman A, et al. Oversight of human participants research: identifying problems to evaluate reform proposals. Ann Intern Med. 2004; 141:282–91. [PubMed: 15313744]
34. Menikoff J, Kaneshiro J, Pritchard I. The Common Rule, updated. N Engl J Med. 2017; 376:613–15. [PubMed: 28103146]

Research directions

- How do parents understand the risks of research compared to the risks of unstudied treatments?
- Would a different approach to informed consent lead to higher or lower rates of research participation?
- What is the proper balance between regulatory protections and parents' rights to enroll their children in research?