PUBLICCITIZEN
1600 20th Street, NW • Washington, D.C. 20009 • 202/588-1000 • www.citizen.org

January 27, 2014

The Honorable Kathleen Sebelius
Secretary
Department of Health and Human Services
200 Independence Ave. SW
Washington, DC 20201

RE: The Surfactant Positive Airway Pressure and Pulse Oximetry Trial in Extremely Low Birth Weight Infants: The SUPPORT Trial – Inadequate Safety Monitoring

Dear Secretary Sebelius:

We are writing in follow-up to Public Citizen's April 10, 2013, letter and May 8, 2013, report regarding the SUPPORT study funded by the National Institutes of Health (NIH) and conducted by approximately two dozen academic medical institutions of the Neonatal Research Network.[1,2] That letter and report highlighted important and material factual omissions regarding the purpose, nature, and risks of the research in the consent forms approved by the institutional review boards (IRBs) and signed by parents of infants enrolled in the SUPPORT study, and also brought to light deficiencies in the study design that resulted in a failure to ensure that risks to subjects were minimized.

To date, the Department of Health and Human Services' (HHS's) response to the serious ethical lapses in the conduct of the SUPPORT study has been unsatisfactory. Rather than taking substantive steps to remedy these ethical lapses, HHS bowed to pressure from NIH and an academic research establishment dependent on NIH for support and stifled appropriate compliance oversight enforcement action by the Office for Human Research Protections (OHRP).

We write to you now to highlight the following additional important issues related to the SUPPORT study that have come to our attention and also have not been adequately addressed by HHS:

(1) The SUPPORT study protocol appears to have lacked a safety monitoring plan for *separately* monitoring for differences in severe retinopathy of prematurity (ROP) between the two experimental oxygen study groups. If such a plan had been

[1] Carome MA, Wolfe SM. Letter to Secretary of Health and Human Services Kathleen Sebelius regarding the SUPPORT study. April 10, 2013. http://www.citizen.org/documents/2111.pdf. Accessed January 20, 2014.
[2] Carome M, Wolfe S, Macklin R. Analysis of the complete protocol and consent form for the SUPPORT study: Lack of informed consent and a failure to ensure that risks were minimized. May 8, 2013. http://www.citizen.org/documents/2124.pdf. Accessed January 20, 2014.

implemented, the study likely would have been terminated early, sparing some extremely premature infants enrolled in the study from suffering severe ROP or death.

(2) In spite of evidence we presented previously,[3] the SUPPORT study investigators — in an attempt to defend the adequacy of their study consent forms — have continued to repeatedly assert that they all had no expectation that the low-oxygen group subjects would have a higher mortality rate than the high-oxygen group subjects and indeed were surprised when the final study results revealed such an outcome. We provide below additional clear evidence that before the study began, there was an awareness among at least some of the investigators — including among the lead investigators who developed the study protocol — that an increased mortality rate in the low-oxygen group was one foreseeable, plausible outcome of their study. Moreover, as neonatology experts, the SUPPORT study investigators had an ethical obligation to thoroughly research the literature, to understand areas of ongoing uncertainty regarding oxygen management, and to be aware of all plausible study risks. To not have known that an increased mortality rate in the low-oxygen group was one foreseeable, plausible outcome of their study would have constituted reckless ignorance.

(3) The SUPPORT study was one of five concurrently planned, coordinated, and conducted studies around the world using nearly identical study designs for testing low- versus high-oxygen interventions in extremely premature infants. To minimize risks to subjects across all five studies, the protocols should have specified a mechanism for joint safety monitoring. Ideally, there should have been a formal data and safety monitoring plan involving interim analyses of pooled data from all five studies combined. At a minimum, there should have been a plan for informally sharing any troubling safety signal arising in one study with the investigators and data monitoring committees for the other studies. No such formal or informal plan was described in the SUPPORT protocol. Of note, the members of the data monitoring committee for at least one of the five studies recognized early on the need to see interim data from the other four parallel studies to ensure the safety of subjects across all trials, but their attempts to obtain data from the other studies were either ignored or rebuffed.

We describe below each of these issues in detail and conclude by asking key questions for which the parents of SUPPORT study subjects and the public deserve clear answers from HHS.

A. Background

As you are aware, the SUPPORT study involved two simultaneous complex experiments. In one experiment, the babies were randomly divided into two groups that each received a different treatment to assist their breathing (ventilation of the lungs) following delivery.[4]

For the other experiment (the oxygen experiment), babies assigned to each of the two ventilation groups were further randomly divided between a low-oxygen group and a high-oxygen group.[5]

[3] *Ibid.*

[4] SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Early CPAP versus surfactant in extremely premature infants. *N Engl J Med.* 2010;362(21):1970-1979.

For the low-oxygen group, the SUPPORT study investigators tried to maintain the babies' blood oxygen levels in a low target range (oxygen saturation level of 85 to 89 percent), and for the high-oxygen group in a high target range (oxygen saturation level of 91 to 95 percent), regardless of the infants' clinical status.

The primary efficacy outcome measure for the oxygen experiment was a combination of severe retinopathy of prematurity (ROP, which can lead to visual impairment and blindness and often requires surgery to preserve vision), death before discharge from the hospital, or both.

For both experimental oxygen groups, oxygen monitors relied upon by the medical teams caring for the infants in the study displayed either intentionally falsely high (low-oxygen group) or intentionally falsely low (high-oxygen group) values when the infants' actual oxygen saturation levels were between 85 and 95 percent.

Premature infants enrolled in the SUPPORT study were not given the same care with respect to oxygen management that otherwise similar infants would have received at the same participating hospitals. In particular, oxygen management of enrolled infants lacked the following features of usual care:

(1) fully functional, properly operating pulse oximeters that displayed accurate oxygen saturations for use by health care providers to guide care;

(2) access to the entire range of target oxygen saturations endorsed in guidelines (85 to 95 percent) for management of premature infants, including the possibility of employing the center of this range (88 to 92 percent); and

(3) adjustment of supplemental oxygen and oxygen saturation targets based on an assessment of risks and benefits for each infant's particular characteristics. Some of the clinical factors that are often considered in individualizing oxygen management include level of prematurity; capillary refill time (a simple physical exam test to assess the adequacy of tissue perfusion); cardiopulmonary, hepatic, and renal function; hematocrit; intravascular volume; acid-base status; oxygen requirements in the toxic range; and clinical signs suggestive of impending necrotizing enterocolitis.

For the low-oxygen group infants, the experimental strategy for attempting to maintain oxygen saturation levels at 85 to 89 percent using oxygen monitors that displayed falsely high readings predictably caused, on average, lower levels of oxygen exposure than would have occurred under usual care using accurately reading oxygen monitors. The low-oxygen intervention presented the foreseeable risks of neurologic injury and death.

In contrast, for the high-oxygen group infants, the experimental strategy for attempting to maintain oxygen saturation levels at 91 to 95 percent using oxygen monitors that displayed falsely *low* readings predictably caused, on average, higher levels of oxygen exposure than

[5] SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med.* 2010;362(21):1959-1969.

would have occurred under usual care using accurately reading oxygen monitors. The high-oxygen intervention presented the foreseeable risk of severe ROP.

B. Inadequate safety monitoring plan: Apparent failure to monitor for severe ROP as a separate adverse event and to terminate the study early because of harm to subjects in the high-oxygen group

Minimization of risks to research subjects requires adequate safety monitoring. Both death and severe ROP comprised the primary risks of the SUPPORT study's oxygen experiment, and each should have been monitored *separately as adverse events* during the conduct of the trial in order to minimize risks to subjects. If separate monitoring of both had been implemented, the study likely would have been terminated early, sparing some extremely premature infants enrolled in the study from suffering severe ROP or death.

However, as reflected in the following excerpts from the data and safety monitoring plan in the SUPPORT study protocol, it appears that unlike death, severe ROP was not monitored *separately as an adverse event* during the course of the trial. Instead, severe ROP apparently was monitored only in combination with death as a component of the composite primary efficacy endpoint, and the study was not terminated early. The protocol stated:[6]

> **4.4 Adverse Events**
>
> Serious and unanticipated adverse events may be anticipated in this vulnerable population. Data on the following potential adverse events that may be related to the study maneuver will be recorded and evaluated as part of continuous safety monitoring during the trial by RTI:
>
> 1. Air leak on admission to the NICU, or during the initial 14 days of life
> 2. The need for chest compressions, and/or epinephrine in the delivery room or NICU
> 3. The occurrence of severe IVH (Grades 3-4, Papile)
> 4. Death
>
> These outcomes will be evaluated on a monthly basis by RTI [Research Triangle Institute], and if the incidence of any of these outcomes is determined to be 5% -10% greater in any arm of the study, this information will be provided to the Study PI and committee and the DSMC for immediate consideration, and evaluated for consideration of termination of the study or treatment arm.
>
> **4.5 Data Safety Monitoring Committee**
>
> The Data Safety Monitoring Committee will review the progress of the study with respect to efficacy and adverse events in a sequential fashion by using interim monitoring boundaries. Obrien-Fleming boundaries will be used for efficacy monitoring and will be constructed for four looks at the data at 25%, 50%, 75%, and 100% of outcome

[6] NICHD Neonatal Research Network. The surfactant positive airway pressure and pulse oximetry trial in extremely low birth weight infants: The SUPPORT trial. August 28, 2004 (revised September 16, 2004; updated March 28, 2005). http://www.nih.gov/icd/od/foia/library/Protocol.pdf. Accessed January 20, 2014.

assessment. Pocock boundaries will be used for adverse event monitoring and will be viewed after every 30 infants have been enrolled. A special adverse event form will collect the information so that it may be entered into the data base in a timely fashion.

Monitoring of the *composite* primary efficacy endpoint of severe ROP or death before discharge during the conduct of the SUPPORT study as designed was not sufficient for monitoring safety related to the occurrence of severe ROP because of the following factors:

(1) The SUPPORT study's oxygen experiment involved only two experimental groups (the low-oxygen group and the high-oxygen group) and no usual care (or current-practice) control group; and

(2) The two components of the composite primary efficacy endpoint — death and severe ROP — were countervailing, but asymmetric, potential harms:

(a) For the high-oxygen group infants, the experimental strategy for attempting to maintain oxygen saturation levels at 91 to 95 percent using oxygen monitors that displayed falsely low readings predictably caused, on average, higher levels of oxygen exposure than would have occurred under usual care using accurately reading oxygen monitors. As a result, the research procedures for these subjects presented a reasonably foreseeable increased risk of suffering severe ROP (a risk that was not described in 20 of 22 IRB-approved SUPPORT study consent forms as required by HHS human subjects protection regulations at 45 CFR 46.116(a)(2)).[7]

(b) For the low-oxygen group infants, the experimental strategy for attempting to maintain oxygen saturation levels at 85 to 89 percent using oxygen monitors that displayed falsely high readings predictably caused, on average, lower levels of oxygen exposure than would have occurred under usual care using accurately reading oxygen monitors. As a result, the research procedures for these subjects presented a reasonably foreseeable risk of death (a risk that was not disclosed in any of 22 IRB-approved SUPPORT study consent forms).

Adequate safety monitoring of the study as designed would have required periodic checking for differences between the low-oxygen and high-oxygen groups for *both* death and retinopathy *separately*. The importance of such separate comparisons was reflected in the way the results were presented in the published paper describing the primary results of the study (see Table 1 below).[8]

[7] IRB-approved consent form for the SUPPORT trial. http://www.citizen.org/documents/support-study-consent-form.pdf. Accessed January 20, 2014.

[8] SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med.* 2010;362(21):1959-1969.

Table 1: Key Major Outcomes from SUPPORT Study

Outcome	Low-Oxygen Group *no./total no.(%)*	High-Oxygen Group *no./total no.(%)*	Adjusted Relative Risk (95% Confidence Interval)	P-value
Primary efficacy outcome: severe ROP or death before discharge	171/605 (28.3%)	198/616 (32.1%)	0.90 (0.76-1.06)	0.20
Severe ROP	41/475 (8.6%)	91/509 (17.9%)	0.52 (0.37-0.73)	<0.001
Death before discharge	130/654 (19.9%)	107/662 (16.2%)	1.27 (1.01-1.60)	0.04

Because severe ROP often requires surgery and can lead to blindness, it represented a clear potential harm to the high-oxygen group infants of significant-enough degree to require separate safety monitoring, as was done for death. Yet the protocol's data and safety monitoring plan did not indicate that it was separately monitored. Indeed, the problem of ROP in premature infants was considered such a serious health problem in premature infants that NIH spent more than $20 million on the SUPPORT study to find out whether using the low-oxygen intervention would reduce the incidence of this important adverse outcome in comparison to the high-oxygen intervention without causing an increase in mortality or brain injury.

In what obviously came as no surprise to the SUPPORT study investigators and was not an unexpected finding, the study results, after SUPPORT was concluded, demonstrated that the high-oxygen group babies had a highly significant increase in retinopathy in comparison to the low-oxygen group babies (17.9 percent versus 8.6 percent, respectively; $p<0.001$).[9]

If the incidence of severe ROP had been monitored separately as an important adverse event in the high-oxygen group at increased risk for this adverse outcome and compared to the incidence in the low-oxygen group, the trial conceivably could have been stopped early, thus preventing the occurrence of retinopathy and avoiding retinal surgery in many high-oxygen group infants.

To estimate the approximate point at which one of the planned interim analyses of study data would have demonstrated a statistically significant difference in the rate of severe ROP between the high- and low-oxygen groups, we performed a series of hypothetical interim analyses using the chi-square test based on enrollments of 25, 50, and 75 percent of the actual final enrollment (i.e., 1,316 infants). Because interim data were not available to us for these analyses, we assumed that the following factors remained constant throughout the study: the incidence of severe ROP, the proportion of high- and low-oxygen group subjects, and the proportion of subjects in each group who survived to the time of discharge and had a determination made regarding whether severe ROP had developed.

[9] *Ibid.*

The contingency tables below (Tables 2-4) summarize our analysis. Following each table is the chi-square statistic without Yates correction and two-tailed P-value (uncorrected for multiple looks):

Table 2: 25 Percent of Final Enrollment*Group	**No Severe ROP**	**Severe ROP**	**Total**
High-Oxygen	104	23	127
Low-Oxygen	109	10	119
Total	213	33	246

*Assumes 329 subjects enrolled and 246 survived to discharge and had retinopathy status determined
χ^2=4.984; P=0.0256

Table 3: 50 Percent of Final Enrollment*

Group	**No Severe ROP**	**Severe ROP**	**Total**
High-Oxygen	209	45	254
Low-Oxygen	217	21	238
Total	426	66	492

***Assumes 658 subjects enrolled and 492 survived to discharge and had retinopathy status determined
χ^2=8.366; P=0.0038

Table 4: 75 Percent of Final Enrollment*

Group	**No Severe ROP**	**Severe ROP**	**Total**
High-Oxygen	314	68	382
Low-Oxygen	325	31	356
Total	639	99	738

*Assumes 987 subjects enrolled and 738 survived to discharge and had retinopathy status determined
χ^2=13.118; P=0.0003

Based on the above analyses, with respect to the harmful outcome of severe ROP, a statistically significant greater incidence of harm in the high-oxygen group might have been detected after reaching just 25 percent of target subject enrollment, and almost certainly would have been detected after reaching 50 percent of target enrollment. Early termination of the study following enrollment of either one-quarter or one-half of the projected final target enrollment likely would have spared some of the subsequently enrolled infants randomized to the high-oxygen group from developing severe ROP that resulted from receiving a higher level of oxygen exposure than they would have otherwise received if they had not been enrolled in the study. We are not able to reliably estimate the number of children who would have been spared severe ROP had the study been terminated early because the study lacked a current-practice control group.

It is important to recognize that the inclusion of a plan to separately monitor for severe ROP as an important adverse outcome during the conduct of the study would not have been sufficient to

address the other fundamental flaw in the SUPPORT study design — the lack of a usual-care control group. By experimentally increasing oxygen exposure in one study group relative to current practice and lowering it in the other, the harms resulting from each experimental intervention *relative to current practice* could not be monitored for or determined, and therefore, risks to subjects were not minimized.

The SUPPORT study investigators may argue that separately monitoring for severe ROP as an adverse event could have led to premature termination of the study and prevented the detection of the higher mortality rate that was seen in the low-oxygen group, which in turn could have led to a dangerous recommendation to routinely target oxygen saturation levels at 85 to 89 percent in all extremely premature infants. (Indeed, as discussed in the next section of our letter, this thought process may explain why separate monitoring for severe ROP did not occur.) However, if the study had been stopped early based on an interim analysis showing a statistically significant higher incidence of severe ROP in the high-oxygen group, without a usual-care control group there would have been no sound basis for concluding that the lower-oxygen intervention was better than usual care. The higher incidence of severe ROP in the low-oxygen group may have been due to oxygen exposure being experimentally raised in the high-oxygen group (relative to usual care), experimentally lowered in the low oxygen group (relative to usual care), or both. The study results were ultimately uninformative in this regard given the lack of a usual-care control group. In addition, interim analyses may have revealed a non-statistically significant higher death rate in the low-oxygen group compared to the high-oxygen group, which should have precluded anyone from making recommendations to modify current practice to routinely using the lower oxygen saturation target in the clinical care of extremely premature infants.

All of these problems could have been avoided with an alternative study design that employed a usual-care control group and low-oxygen experimental group. Such a design could have informed the medical community whether oxygen could be safely lowered to prevent severe ROP without increasing the mortality rate. Moreover, because experimentally lowering oxygen was unlikely to increase the incidence of severe ROP and blindness relative to current practice, it would not have been necessary to monitor this outcome as a separate adverse event.

C. The higher mortality rate in low-oxygen group infants: *Not* a surprising finding

As noted above, for the low-oxygen group infants, the experimental strategy for attempting to maintain oxygen saturation levels at 85 to 89 percent using oxygen monitors that displayed falsely high readings predictably caused, on average, lower levels of oxygen exposure than would have occurred under usual care using accurately reading oxygen monitors.

Over the past several months, in an awkward attempt to explain why death was not identified as a risk of the research in the SUPPORT study consent forms, the study investigators have repeatedly asserted that they all had no expectation that the low-oxygen group subjects would have a higher mortality rate than the high-oxygen group subjects, and indeed some investigators

have indicated that they were surprised when the final study results revealed such an outcome.[10,11,12,13]

In contrast, the investigators in recent months have not asserted that they were surprised to find a higher rate of severe ROP in the high-oxygen group than in the low-oxygen group, and as previously noted, such a finding undoubtedly was not a surprise.

All physicians understand the well-established pathophysiologic relationship in which increasing degrees of hypoxemia result in progressively higher mortality rates. At the time of the SUPPORT study, all neonatologists knew that increasing the degree of hypoxemia in premature infants *at some point* would increase the infants' mortality rate. The exact shape of the curve for the relationship between the degree of hypoxemia and mortality and the exact threshold of oxygen exposure below which mortality will start to increase in premature infants were not known at the time the SUPPORT study was conducted and remain unknown today.

The SUPPORT study investigators may have believed that targeting oxygen saturations at 85 to 89 percent in extremely premature infants was unlikely to cross below the oxygen exposure threshold that would increase the mortality rate. However, they did not know for certain that this was the case, and the available data from the medical literature cited by the investigators in their protocol were clearly insufficient to prove that maintaining oxygen saturations at 85 to 89 percent would *not* have an adverse impact on the mortality rate of extremely premature infants. Indeed, assessing whether the lower oxygen saturation target could decrease the incidence of severe ROP in severely premature infants without increasing mortality was one of the major reasons for conducting the study.

As neonatology experts, the SUPPORT study investigators had an ethical obligation to thoroughly research the literature, to understand areas of ongoing uncertainty regarding oxygen management, and to be aware of all plausible study risks. To not have known that an increased mortality rate in the low-oxygen group was one foreseeable, plausible outcome of their study would have constituted reckless ignorance.

Investigators in New Zealand undertaking a similarly designed study (discussed below) understood that regardless of their expectations about the study outcome, it was appropriate to

---

[10] Transcript for the Diane Rehm Show from WAMU and National Public Radio: Clinical trials and premature babies. April 17, 2013. http://thedianerehmshow.org/shows/2013-04-17/clinical-trials-and-premature-babies/transcript. Accessed January 20, 2014. (SUPPORT study investigator Dr. Edward Bell stated, "As a matter of fact, there was no reason for us to suspect that there would be a difference in mortality, and that was a surprising finding.")

[11] Finer NN, Bell EF, Van Meurs K. Consent forms in a clinical trial of premature babies (letter to the editor). *The New York Times*. April 18, 2013. http://www.nytimes.com/2013/04/19/opinion/consent-forms-in-a-clinical-trial-of-premature-babies.html. Accessed January 20, 2014. ("When the study was planned, the best evidence showed that lower oxygen targets — even lower than used in the study — resulted in less eye disease without a higher death rate. The finding of a higher death rate in one study group was not anticipated.")

[12] Carlo WA, Bell EF, Walsh MC. Oxygen saturation targets in extremely premature infants (letter to the editor). *N Engl J Med*. 2013;368(20):1949-1950. ("Death was included in the primary outcome because it competes with retinopathy, not because a difference in mortality was expected.")

[13] D'Angio C, Finer N, Newman N, et al. Misconceptions about the SUPPORT trial. Posted September 10, 2013. http://www.regulations.gov/#!documentDetail;D=HHS-OPHS-2013-0004-0067. Accessed January 20, 2014.

explain to subjects' parents in the consent forms the well-known and previously established separate risks of death from lowering oxygen exposure, as well as ROP from raising oxygen exposure, for premature infants enrolled in the study. It is unclear why the SUPPORT study investigators in the U.S. would have lacked such an understanding.

Although the SUPPORT study investigators undoubtedly were *hopeful* at the onset of the study that the low-oxygen intervention would result in a decreased incidence in severe ROP without a concomitant increase in mortality, there is clear evidence that before initiating the study at least some of them were aware that a higher mortality rate in the low-oxygen group relative to the high-oxygen group could have been one plausible finding of their study. In particular, the SUPPORT study investigators cited in their protocol a 2003 paper by Cole et al, published in the journal *Pediatrics,* discussing the planning and design of studies comparing the low- and high-oxygen interventions that were to be used in their study[14] (see reference 55 in the SUPPORT study protocol[15]). The authors of this paper were members of the Pulse Oximetry Saturation Trial for Prevention of Retinopathy of Prematurity (POST ROP) Planning Study Group.

The POST ROP study was to be a multicenter, multinational prospective trial to evaluate different levels of oxygen in premature babies. It is our understanding that the POST ROP study comprised the Benefits of Oxygen Saturation Targeting (BOOST) II studies in the United Kingdom (UK), Australia, and New Zealand, and the Canadian-funded Canadian Oxygen Trial (COT). The SUPPORT study protocol made reference to this study as follows: "The oxygen saturation monitoring portion of our study will be designed to parallel the planned POST-ROP trial…"

Of note, Dr. Waldemar Carlo at the University of Alabama at Birmingham — who was one of the lead investigators for the SUPPORT study and was on the Neonatal Research Network working group that developed and wrote the SUPPORT study protocol — was a member of the POST ROP Planning Study Group. The acknowledgements section at the end of the Cole et al paper states that Dr. Carlo was among the POST ROP Planning Study Group members who reviewed and critiqued the paper.[16]

The 2003 Cole et al paper makes clear that when the SUPPORT study and the parallel POST ROP studies were being designed, there were real concerns among neonatologists that the low-oxygen intervention could expose extremely premature infants to an increased risk of death or neurologic injury. In particular, the paper emphasized that large numbers of patients would have to be studied to address concerns about mortality risk with the low oxygen dose, noting the following:[17]

---

[14] Cole CH, Wright KW, Tarnow-Mordi W, Phelps DL. Resolving our uncertainty about oxygen therapy. *Pediatrics.* 2003;112:1415-1419.
[15] NICHD Neonatal Research Network. The surfactant positive airway pressure and pulse oximetry trial in extremely low birth weight infants: The SUPPORT trial. August 28, 2004 (revised September 16, 2004; updated March 28, 2005). http://www.nih.gov/icd/od/foia/library/Protocol.pdf. Accessed January 20, 2014.
[16] Cole CH, Wright KW, Tarnow-Mordi W, Phelps DL. Resolving our uncertainty about oxygen therapy. *Pediatrics.* 2003;112:1415-1419.
[17] *Ibid.*

> Several hundred patients (15-25 centers) may be sufficient to demonstrate important differences in severe ROP. **However, a much larger sample (and many more collaborators) will be needed to exclude smaller, important differences in outcomes such as mortality and disability to adequately address real concerns about the safety of lower oxygen tensions. For example, a 5% difference in an outcome of death or cerebral palsy is "small" but would have major implications for public health. Preliminary calculations suggest that the trial may require a sample size between 2000 and 4000 extremely low gestational age infants (born at <28 weeks' gestation) to answer these important questions.** Participation of centers that undertake long-term follow-up in >90% of their survivors will be necessary. [Emphasis added]

Emphasizing the need for a data and safety monitoring committee and plan for the POST ROP studies, the 2003 *Pediatrics* paper stated the following:[18]

> It is also essential, both ethically and scientifically, to have an external monitoring committee to ensure that if **major differences between the groups with respect to outcomes such as death or severe ROP are detected, they will be detected during the recruitment phase.** Appropriate decisions regarding study termination or continuation can be achieved if stringent stopping rules for the Data Monitoring and Safety Committee are based on evidence beyond reasonable doubt of net clinical benefit or harm or futility of finding a difference before recommending trial termination. Evidence of net benefit or harm from one outcome should be considered in the context of other major outcomes. For example, **it would be inappropriate to terminate recruitment because of a 3% reduction in severe ROP in the lower oxygen group before the trial had accumulated sufficient power to exclude a 6% increase in mortality or severe neurodevelopmental impairment in the same group**. In this case, if the trial were terminated prematurely and lower oxygen became the clinical standard, for every infant whose sight was saved, 2 would die or survive with major disability. [Emphasis added]

Further indication of the serious concern among some neonatologists that the low-oxygen intervention could expose extremely premature infants to an increased risk of death or neurologic injury was provided in the 2005 version of the consent form used in the New Zealand BOOST II study discussed above, which included the following:[19]

> Too low oxygen in the blood for long periods may 1) increase the risk the baby will not survive or contribute to poor growth; 2) raise blood pressure in the lungs and contribute to bronchopulmonary dysplasia; 3) damage the brain cells and lead to developmental problems. . . . The aim of this study is to determine, within the range of oxygen saturation values currently used in the treatment of preterm babies (85-95%), **whether targeting the lower end of this range (85-89%) compared to upper end of this range (91-95%), beginning within 24 hours of birth, is safe and effective in reducing serious vision**

[18] *Ibid*.

[19] Lantos JD. OHRP and Public Citizen are wrong about neonatal research on oxygen therapy research. The Hastings Center Bioethics Forum. Posted April 18, 2013. http://www.thehastingscenter.org/Bioethicsforum/Post.aspx?id=6306&blog+140. Accessed January 20, 2014. Cited as "BOOST-NZ consent form, July 2005, personal communication from Brian Darlow, principal investigator of BOOST-NZ."

> **(ROP) and lung (BPD) problems without increasing mortality or neurodevelopmental disability**. [Emphasis added]

Finally, members of the data monitoring committee for the BOOST II UK study noted the following in a recent commentary article discussing the four POST-ROP studies and the SUPPORT study: [20]

> Evidence from controlled trials had shown that if the oxygen levels are relatively high there is an increased risk of the infant developing the blinding condition retinopathy of prematurity. On the other hand, observational data had suggested that keeping levels of arterial oxygen relatively low might result in increased mortality, and neurological handicap among long term survivors.
>
> During the past decade, five similar trials - in the USA, New Zealand, Canada, Australia and the UK - **were organized more or less concurrently to investigate this therapeutic dilemma**. Although they were separately organized and funded, all of them compared different oxygen tension targeting strategies intended to minimize both mortality and serious morbidity. **It was recognized at the outset that none of the five trials, individually, would have the statistical power to provide a reliable estimate of survival without serious morbidity 18 to 24 months after birth.** Indeed, two of these trials were only funded on the understanding that the data from several similar trials would be combined. [Emphasis added]

All of the above statements could not be clearer. At least some of the expert neonatologists involved in the design of the SUPPORT study and parallel studies to be conducted in other countries were well aware when designing these studies that there were real concerns within the neonatology community that the low-oxygen intervention to be used in these studies might increase the risk of death.

In view of the commentary paper authored by members of the POST ROP Planning Study Group and cited by the SUPPORT study investigators in their own protocol, it is remarkably disingenuous of the SUPPORT study investigators to now assert that because some of them were surprised to find a higher mortality rate in the low-oxygen group subjects in their study, it was not necessary to inform parents about the reasonably foreseeable risk of death for subjects assigned to the low-oxygen group. Such a finding was one reasonably foreseeable and highly plausible outcome. This undoubtedly is one of the reasons why, as noted above, death was to be monitored according to the SUPPORT study protocol and was a component of the primary outcome being studied.

The investigators certainly had hoped for a different result and perhaps had reason to be disappointed when they found the significantly higher mortality rate in the low-oxygen group, but they could not genuinely have been shocked by the fact that when they lowered oxygen exposure, the mortality rate increased in premature infants in the study. Indeed, none of the

[20] Chalmers I, Altman DG, McHaffie H, et al. Data sharing among data monitoring committees and responsibilities to patients and science. *Trials.* 2013;14:102. http://www.trialsjournal.com/content/14/1/102. Accessed January 20, 2014.

investigators at any point in discussing the results of the SUPPORT study have proposed an alternative hypothesis to explain the higher mortality rate in the low-oxygen group compared to the high-oxygen group. They implicitly recognize that the well-established pathophysiologic relationship in which decreasing oxygen exposure results in increasing hypoxemia and mortality is undoubtedly the explanation for their study findings, whether unexpected by some of the investigators or not.

Finally, we strongly suspect that the apparent failure to monitor separately for severe ROP and to have stopping criteria based on finding a difference in the incidence of severe ROP between the high- and low-oxygen groups at the time of any planned interim analyses was due to concern among the investigators that stopping the oxygen experiment early based on a statistically significant difference in ROP between groups could have resulted in a failure to detect a difference in the mortality rate between the two study groups.

D. Failure to minimize risks to subjects by not having a data and safety monitoring plan involving interim analyses of pooled data from the SUPPORT study and the four concurrent POST ROP studies

As noted above, the SUPPORT study and the four POST ROP studies were five parallel studies concurrently planned and conducted around the world using nearly identical study designs comparing high- and low-oxygen interventions in extremely premature infants. The planning and designing of these studies were coordinated,[21,22] and the investigators for each study obviously were aware of the other studies.

To minimize risks to subjects across all five studies, the protocols should have specified a mechanism for joint safety monitoring. Ideally, there should have been a formal data and safety monitoring plan involving interim analyses of pooled data from all five studies combined. At a minimum, there should have been a plan for informally sharing any troubling safety signal arising in one study with the investigators and data monitoring committees for the other studies. No such formal or informal plan was described in the SUPPORT protocol. Instead, monitoring of data and safety was conducted independently and separately across these studies. Not having pooled monitoring across all five studies, even in some informal manner, represents a troubling failure to ensure the protection of human subjects.

Disturbingly, members of the data monitoring committee for the BOOST II UK study recognized the importance of seeing interim data from the other four parallel studies, but their attempts to obtain data from the other studies, which began as early as 2006, apparently were ignored or rebuffed by the data monitoring committees for the other four parallel studies.[23] For example,

---

[21] Cole CH, Wright KW, Tarnow-Mordi W, Phelps DL. Resolving our uncertainty about oxygen therapy. *Pediatrics.* 2003;112:1415-1419.

[22] NICHD Neonatal Research Network. The surfactant positive airway pressure and pulse oximetry trial in extremely low birth weight infants: The SUPPORT trial. August 28, 2004 (revised September 16, 2004; updated March 28, 2005). http://www.nih.gov/icd/od/foia/library/Protocol.pdf. Accessed January 20, 2014.

[23] Chalmers I, Altman DG, McHaffie H, et al. Data sharing among data monitoring committees and responsibilities to patients and science. *Trials.* 2013;14:102. http://www.trialsjournal.com/content/14/1/102. Accessed January 20, 2014.

they noted the following in a recent commentary article discussing the four POST-ROP studies and the SUPPORT study:

> Accordingly, because data from the other four trials were of obvious relevance to our responsibility, the chair of our DMC wrote to the DMC chairs of the other trials in 2006, expressing the **'hope that we can help each other fulfil our respective commitments to the babies being treated in these trials'** (Emails sent 11 and 15 July 2006).
>
> **No response was received from the other DMC chairs for several years**; but consideration of the proposal became urgent when, more than three years later, in 2009, the management group of the US trial sent results, in advance of publication, to those associated with the trials that were still recruiting.

E. Conclusions and requested actions

For each of the three critical issues described above, parents of subjects enrolled in the SUPPORT study, as well as the public, deserve clear answers to the following questions:

(1) With respect to monitoring separately for difference in the incidence of severe ROP between groups:

(a) Was ROP monitored separately during the course of the trial as an important adverse event? If not, why not?

(b) If ROP was monitored separately during the course of the trial, at the time of any of the planned interim analyses, did the difference in the incidence of severe ROP between the low- and high-oxygen groups reach statistical significance? If so, when did this occur, and why wasn't the study terminated at that point?

(c) Did OHRP consider the lack of appropriate safety monitoring during the SUPPORT study and the lack of a usual-care control group when it evaluated adequacy of the SUPPORT study design?

(d) Since the SUPPORT study investigators must have recognized that the incidence of severe ROP was likely to be higher in the high-oxygen group and have voiced no surprise in finding this result, does HHS agree or disagree with OHRP's finding that the IRB-approved consent forms failed to comply with HHS regulations under 45 C.F.R. 46.116(a)(2) by not disclosing severe ROP as a risk of the research?

(2) With respect to the investigators' statements about being surprised to find a higher mortality rate in the low-oxygen group: Since the SUPPORT study investigators either knew or should have known prior to initiating the study that an increased death rate in the low-oxygen group was a foreseeably plausible outcome of the SUPPORT study, does HHS agree or disagree with OHRP's finding that the IRB-approved consent forms failed to comply with HHS regulations under 45 C.F.R. 46.116(a)(2) by not disclosing death as a risk of the research?

(3) With respect to the failure to establish a plan to monitor data pooled across the SUPPORT study and the four POST ROP studies:

(a) Why didn't the SUPPORT study protocol include a plan for joint safety monitoring with the POST ROP studies, either formally, via pooled interim analyses across all five studies, or informally, by sharing any troubling safety signals arising in one study with the investigators and data monitoring committees for the other studies?

(b) Were the SUPPORT study investigators or NIH officials aware of the BOOST II UK data monitoring committee's requests for SUPPORT study data for the purposes of pooled interim analysis? If so, what was their response to those requests? Why were the requests not granted?

We urge HHS to provide prompt answers to these important questions. We also request an opportunity to meet with you or your representative to discuss these important issues, which have critically important implications for the safety and welfare of premature infants participating in ongoing clinical trials funded by HHS.

In closing, we renew our April 10, 2013, request that you issue a formal apology to the parents of all 1,316 subjects enrolled in the SUPPPORT study. This apology should be accompanied by a complete divulgence of the previously undisclosed information regarding the nature, purpose, and risks of the research. Such an apology is the most critical step for redressing the ethical lapses that occurred during the conduct of this study.

Please contact us if you have any questions or need additional information.

Sincerely,

Michael A. Carome, M.D.
Director
Public Citizen's Health Research Group

Sidney M. Wolfe, M.D.
Founder and Senior Adviser
Public Citizen's Health Research Group

cc: The Honorable Bill Corr, Deputy Secretary, HHS
The Honorable Howard K. Koh, Assistant Secretary for Health, HHS
Dr. Francis Collins, Director, NIH

Dr. Alan E. Guttmacher, Director, Eunice Kennedy Shriver National Institute of Child Health and Human Development
Dr. Jerry Menikoff, Director, OHRP
Dr. Kristina Borror, Director, Division of Compliance Oversight, OHRP