Ethical Issues In Neonatal/Perinatal Research

The SWAN Conference
Temple, TX
October 6, 2012

Tonse N. K. Raju, MD
Program Director
Pregnancy and Perinatology Branch
Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH
6100 Executive Blvd, Room 4B03
Bethesda, MD, 20892.
Talk Outline
Two Ethical Issues in Perinatal Research

- Ethics of sample size in RCTs
- Ethics of early stopping
Sample Size: What’s Ethics Got to Do With It?

- We strive to enroll “just enough” subjects in RCTs, but, all too often, we enroll either too few subjects, and rarely, we may study too many subjects.

- Both are problematic.
Ethics of Too Small a Sample Size in RCTs

- Increase in Type I and Type II errors with higher risk for wrong conclusions
- Wrong conclusions may harm patients
  - False hope of a cure when the drug does not work (Type I error)
  - False conclusion that the drug is not effective (Type II error) deter future research on the drug
- Participants are unnecessarily studied
- Wasted time, effort, and money
Ethics of Too Large a Sample Size

- Beneficial interventions may be denied to too many subjects
  - How does one feel to be the last control subject in a clinical trial that showed 50% improved outcome?

- Wasted time, resources, efforts, and money
Elements Used in Calculating Sample Size

- Baseline prevalence of the condition
- Anticipated effect size from intervention
- The extent to which we will “tolerate” coming to wrong conclusions
  - Research conclusions are estimates of reality
  - Thus we try to minimize the chances for errors
- If all of us make errors “very rarely,” in the long run, most often we will come to right conclusions.
Tolerance for Error

- Type I error: typically 5% or less (P<0.05)
  - With identical research conditions, we will risk declaring an ineffective drug effective 5% of the time

- Type II error: 0.1 or 0.2; power 90% or 80%
  - With identical research conditions, we will risk declaring an effective drug ineffective between 10 and 20% of the time
Why Researchers Tend to Choose Suboptimal Sample Sizes?

- Small studies can be completed quickly
  - Another scientific paper
  - Promotion, fame
- Less expensive: to appeal to the funding agencies
- In risky interventions, minimize exposing too many subjects to the risks
- “Convenient” sample size: they don’t have enough subjects to enroll.
Ways to “Statistically” Justify Small Sample Sizes?

- Propose a large treatment-effect size: Expect a dramatic improvement from intervention
- Choose lower power (higher Type II error)
- Calculate power “after the study”
Ethical Issues

- Anticipating a large treatment effect (small sample size), and higher Type II error
  - But, smaller effect may be clinically important, or useful to a given person, or to the society
  - By declaring the drug is not effective, others may not undertake such studies
  - It may be years before the same drug is tested again

- Recalculating sample size after the study: problematic: better be transparent than deceptive!
Early or Delayed Enteral Feeding for Preterm Growth-Restricted Infants: A Randomized Trial

**Results**

Full enteral feeding achieved in early feeding group 18 days in late feeding group 21 days $p<0.03$

NEC in early feeding group: 36 (18%) NEC in late feeding group: 30 (15%) $p = 0.42$

**Sample Size**

Unpublished nutrition data from a UK regional database of very low birth weight infants revealed an SD of 9 days in the time taken to reach full enteral feeding. We calculated that 380 infants would be required to show a difference of 3 days in this outcome with 90% power. The incidence of NEC from published literature is ~15% in this population, and a sample of 400 would be sufficient to show a 50% change in the incidence of NEC with 60% power.
How Much ‘Better’ Is Good Enough?

The Magnitude of Treatment Effect in Clinical Trials

Tonse N. K. Raju, MD; Patricia Langenberg, PhD; Ashish Sen, PhD; Otto Aldana, MD

- Median predicted effect size (benefit from surfactant treatment) was 50% (range 15-90%)
- Median effect size detected in the 17 trials was 35% improvement (range 75% to -5%)
- Researchers tend to overestimate the anticipated effect size from their interventions

AJDC 1992; 146: 407-411
When Can We Justify a Large Treatment Effect While Calculating the Sample Size?

In very rare cases $N=1$ may be reasonable
Large Treatment Effect
An Example of N=1 Study

- An intervention helped a monkey to recite the opening monologue of Shakespeare’s *Richard III*
- We will accept N=1
- We will not say, “Well, it is just one monkey, singing only one monologue!”
Small Effect May Have a Big Impact

- A cold remedy reduces symptoms by 5%
- An anti-hypertensive drug that reduces the mean diastolic BP by 2 mm of mercury
- An Alzheimer’s disease therapy improves memory by 7%
- All of the above potentially have huge individual benefit, and societal impact
The Effect Size: “How Much Better is Good Enough?”

- The answer is one of context
- The severity and prevalence of the condition in the society, and its “cost” on human life (financial and physical costs)
- The medical and social impact of reducing/improving such outcomes
- Not paying attention to them, and specifying them in the study may be unethical
Take Home Points

- Strive to justify all of the choices we make in deciding on the elements of sample size calculation
- Include each of the choices in manuscripts
- Justify, especially if large effect size is anticipated, or low power is selected
- Be sensitive to the sacrifices our participants make towards our research
Early Stopping of Clinical Trials
Four Reasons for Early Stopping of RCTs

- For adverse effects
- For efficacy
- For futility
- For other reasons
  - Accumulating information & loss of equipoise
  - Lower than predicted recruitment
Ethical Issues of Stopping Clinical Trials Early

- Benefit
  - No brainer: but, what are the risks?

- For side effects
  - What are the risks?

- For slow recruitment
  - A dilemma for funding agency
    - Wasting of public funds?
    - How to deal with current subjects?
# A Ten-Fold Increase in RCTs Stopped Early for Benefit

<table>
<thead>
<tr>
<th>Year of Publication</th>
<th>RCTs Stopped Early/RCTs Indexed in MEDLINE (%)</th>
<th>RCTs Stopped Early/RCTs in Top-Ranked Journals† (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975–1979</td>
<td>1/6574 (0.01)</td>
<td>0/620 (0)</td>
</tr>
<tr>
<td>1980–1984</td>
<td>1/12,653 (0.008)</td>
<td>1/1175 (0.1)</td>
</tr>
<tr>
<td>1985–1989</td>
<td>10/21,807 (0.05)</td>
<td>9/1938 (0.5)</td>
</tr>
<tr>
<td>1990–1994</td>
<td>19/38,712 (0.05)</td>
<td>15/3106 (0.5)</td>
</tr>
<tr>
<td>1995–1999</td>
<td>41/52,060 (0.08)</td>
<td>35/3594 (1.0)</td>
</tr>
<tr>
<td>2000–2004</td>
<td>71/58,537 (0.1)</td>
<td>47/3859 (1.2)</td>
</tr>
</tbody>
</table>

Significance of trend† (p value) <0.0001 <0.0001

*Data are from Montori et al. RCTs = randomized controlled trials. †Top-ranked journals included The New England Journal of Medicine, Journal of the American Medical Association, The Lancet, Annals of Internal Medicine, and the British Medical Journal. ‡Chi-square test with one degree of freedom.

Briel et al, 2012
Early stopping of randomized clinical trials for overt efficacy is problematic

Dirk Bassler\textsuperscript{a}, Victor M. Montori\textsuperscript{b}, Matthias Briel\textsuperscript{c,d}, Paul Glasziou\textsuperscript{e}, Gordon Guyatt\textsuperscript{c,*}

Abstract

**Objective:** To illustrate controversial issues associated with stopping randomized controlled trials (RCTs) early for apparent benefit.

**Study Design and Setting:** The article presents our review of prior relevant work and our research group’s reflections on early stopping.

**Results:** Compelling evidence suggests that trials stopped early for benefit systematically overestimate treatment effects, sometimes by a large amount. Unresolved controversies in trials stopped early for benefit include ethical and statistical problems in the interpretation of results.

**Conclusions:** The best strategy to minimize the problems associated with early stopping of RCTs for benefit is not to stop early. As an alternative, we suggest a threefold approach: a low $P$-value as the threshold for stopping at the time of interim analyses, not to look before a sufficiently large number of events has accrued and continuation of enrollment and follow-up for a further period.

**Keywords:** Randomized controlled trial; Data-monitoring committee; Stopping rule; Interim analysis; Stopping for benefit; STOPIT-2
All Cautious Drivers will Stop for Good Reasons
But, We Should Justify... Take Home Points

- Both the ethical issues and statistical risks should be considered at the design stages.
- All decisions should be disclosed to the subjects, in publications—transparency.
- Plans should have been made how to continue to support the already-recruited study participants.
- Funders should be included in all decisions.
Thank you

- Clinical trials are not easy to design and carry out
- But, the paths to true discoveries are rarely straightforward