Society for Clinical Trials 34th Annual Meeting

Workshop P9
Data Safety Monitoring Boards for Multi-Regional Clinical Trials

Sunday, May 19, 2013
1:00 – 5:00 PM
Gardner Room
Building a Learning Community among Key Stakeholders
Overview

History of MRCT

DMC Workgroup Goals and Programs

This Afternoon’s Session Overview
Allocation of Studies and Patients

SOUNING BOARD

Ethical and Scientific Implications of the Globalization of Clinical Research
Seth W. Glickman, M.D., M.B.A., John G. MacIntosh, M.D., Eric D. Peterson, M.D., M.P.H., Charles B. Cairns, M.D., Robert A. Harrington, M.D., Robert M. Califf, M.D., and Kevin A. Schullman, M.D.

Economic globalization is an important development of the past half-century. Proponents of globalization highlight the benefits of greater economic growth and prosperity. Critics point to the exacerbation of economic disparities and the exploitation of workers, particularly in developing (i.e., low- and middle-income) countries. Pharmaceutical and device companies have embraced globalization as a core component of their business models, generally in the name of clinical industry-sponsored phase 3 clinical trials as of November 2007 for the 20 largest U.S.-based pharmaceutical companies. We found that approximately one third of the trials (157 of 509) are being conducted solely outside the United States and that a majority of study sites (13,521 of 24,206) are outside the United States. Many of these trials are being conducted in developing countries, including the rapidly evolving countries of Eastern Europe and the former Soviet Union (Fig. 1).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>1995 (N = 150)</th>
<th>2005 (N = 150)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of countries represented</td>
<td>33</td>
<td>70</td>
</tr>
<tr>
<td>No. of patients per trial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>215</td>
<td>661</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>75–830</td>
<td>239–1837</td>
</tr>
<tr>
<td>Multinational trials — no. (%)</td>
<td>25 (16.7)</td>
<td>44 (29.3)</td>
</tr>
<tr>
<td>Information reported about location — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locations not reported</td>
<td>59 (39.3)</td>
<td>13 (8.7)</td>
</tr>
<tr>
<td>Only continents reported</td>
<td>5 (3.3)</td>
<td>8 (5.3)</td>
</tr>
<tr>
<td>Only number of countries reported</td>
<td>6 (4.0)</td>
<td>14 (9.3)</td>
</tr>
<tr>
<td>Names of countries reported</td>
<td>79 (52.7)</td>
<td>113 (75.3)</td>
</tr>
<tr>
<td>Enrollment from each country reported†</td>
<td>1 (4.0)</td>
<td>2 (4.5)</td>
</tr>
<tr>
<td>Countries per trial — no. (%)‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>65 (75.6)</td>
<td>94 (72.9)</td>
</tr>
<tr>
<td>2–10</td>
<td>17 (19.8)</td>
<td>20 (15.5)</td>
</tr>
<tr>
<td>11–20</td>
<td>4 (4.7)</td>
<td>5 (3.9)</td>
</tr>
<tr>
<td>&gt;20</td>
<td>0</td>
<td>10 (7.8)</td>
</tr>
<tr>
<td>Regions represented — % of trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>5.0</td>
<td>8.7</td>
</tr>
<tr>
<td>Eastern Europe and Russia</td>
<td>2.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Middle East</td>
<td>1.3</td>
<td>3.5</td>
</tr>
<tr>
<td>Asia</td>
<td>8.8</td>
<td>6.1</td>
</tr>
<tr>
<td>United States</td>
<td>53.8</td>
<td>42.6</td>
</tr>
<tr>
<td>Western Europe</td>
<td>40.0</td>
<td>36.5</td>
</tr>
</tbody>
</table>

* Percentages may not sum to 100 because of rounding.
† The percentages are based on 25 multinational trials in 1995, and 44 in 2005.
‡ The percentages are based on the number of articles that reported country-level information (86 articles in 1995 and 129 articles in 2005).
Rapid movement of studies internationally

Reliance on local Institutional Review Boards (IRB’s) and local or international Contract Research Organizations (CRO’s)

Lack of “transnational” rules and standards applicable to conduct of international trials

Difficulties in finding qualified investigators and tailoring informed consent to local culture and conditions

The process for consensus building on these and other issues

Can we improve communication between International Ethics Committees (ECs), Sponsors, Regulators, and other stakeholders?
The MRCT Center Current Focus Areas

- Investigator Training
- Data and Safety Monitoring
- Ethics Guidance and Support
- Clinical Data Sharing & Transparency
- Global Regulatory Engagement
Multi-Regional Clinical Trials:
Enhancing Respect for Research Participants, Safety and Fairness

A draft report from the Multi-Regional Clinical Trials Project
to be reviewed and discussed at the MRCT Second Summit

January 19-20, 2010
Co-Chairs:

Joe Massaro (BU)    Charles Knirsch (Pfizer)

- Alan Eggleston (CMed)
- Martha Brumfield (CPI)
- Jeff Cooper (Huron)
- Dennis Dixon (NIH)
- Susan Ellenberg (Penn)
- Joan Herbert (MMV)
- Sonali Kochhar (Path)
- John Orloff (Novartis)
- Jerry Sadoff (J & J)
- Steve Snapinn (Amgen)
- Yoko Tanaka (Lilly)
- Janet Wittes (Stats Collaborative) – Lead author of Phase 1 report

- Mark Barnes, Barbara Bierer (MRCT, ad hoc)
Data Safety Monitoring Project

1. Identify qualified DSMB members from the developing world

2. Educate and train DSMB members for trials in the developing world

3. Apprentice DSMB members from emerging markets to serve on boards
Data Safety Monitoring Goal

**Impact:** Increased engagement of experts from emerging world on Data Monitoring Committees for multi-regional trials.

Goal – to identify, train, recruit experts from emerging regions who have expertise in medicine or statistics, experience in clinical trials, and who would like to serve on Data Monitoring Committees.
Data Safety Monitoring Program Plan

1. Identify qualified DSMB members from the developing world
   Fogarty Institute; agreement they would solicit qualified Fogarty International Clinical Research Scholars & Fellows for the program

2. Educate and train DSMB members for trials in the developing world
   Today’s session and partnership with Society of Clinical Trials and co-sponsor a training workshop at the SCT meeting (May 17, 2013, Boston)

1. Apprentice DSMB members from emerging markets to serve on boards
   Pharma members of workgroup are reviewing trials:
   • to be conducted in emerging countries
   • in the pipeline to start in Summer/Fall of 2013
   • would be appropriate to allow fellows to participate

Timeline – 6-12 fellows to be trained in May and start participation in Spring/Fall of 2013
Today’s DSMB Training Curriculum

**Target Audience** - Investigators, ethicists and statisticians who have never served on a DSMB or need a refresher

**SESSION A (1/2 DAY) – Lead by SCT**
- What is the role of the DSMB, composition
- Charter – role, what is it, how used
- Cover the various roles (chair, presenters, etc)
- How to present to the DSMB
- Role-playing based on real trials
- Stopping rules

**SESSION B (1/2 DAY) – Lead by MRCT**
- Provides further depth on issues that arise from global trials via Case studies
Outline

• Multi-Regional Clinical Trials (MRCT)
  – Concept
  – Example
  – Issues that could occur in MRCTs

• Case Study
  – DSMB role in assessing results in MRCTs
Sources of variability in estimates of treatment effect / response and other factors

*Slide courtesy of Robert T. O’Neill, Ph.D., Senior Statistical Advisor to FDA
Ticagrelor vs. Clopidogrel in Patients with Acute Coronary Syndromes (PLATO trial)†

- Multicenter, randomized, double-blind
- Primary endpoint: MI, stroke, death from vascular causes:
  - Overall rate: 9.8% in Ticagrelor, 11.7% in Clopidogrel at 12 months
  - Hazard Ratio: 0.84 (95% CI of 0.77 to 0.92); p<0.001

MRCT Example (Cont’d)

- Significant treatment effect within region NOT necessary to show
- Hope: Consistent treatment effect (e.g., above, desire all HRs < 1)
- Overall event rate hire in Central/South America: Not necessarily a concern; effect size consistent with other OUS regions.
• Inconsistent treatment effect size across regions
  – Benefit of Ticagrelor not seen in US
• According to investigators, “these findings may have been due to chance, given the large number of tests performed”
  – (overall, assessment of treatment effect tested within 66 subgroups)
According to investigators,

- “...raises the questions of whether geographic differences between populations of patients or practice patterns influenced the effects of the randomized treatments, although no apparent explanations have been found.”
But Different Event Rates across Regions May be a Concern Even when Effect Sizes Consistent Across Regions

Suicide Rates in Short-term Randomized Controlled Trials of Newer Antidepressants
Tarek A. Hammad, MD, PhD, MSc, MS, Thomas P. Laughren, MD, and Judith A. Racoosin, MD, MPH

(J Clin Psychopharmacol 2006;26:203–207) *

TABLE 3. Rates of Suicide per 100,000 Person-years and Poisson-based 95% CIs for Active-controlled Depression Trials by Drug Group and Location (11,883 Patients, 94 Trials, 16 Cases)

<table>
<thead>
<tr>
<th>Drug Group</th>
<th>Location</th>
<th>No. Patients</th>
<th>No. Person-years</th>
<th>No. Suicides</th>
<th>Rates/100,000 Person-years</th>
<th>Poisson-based 95% CIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAs</td>
<td>Non–NA</td>
<td>2702</td>
<td>319</td>
<td>6</td>
<td>1881</td>
<td>690–4094</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>1451</td>
<td>176</td>
<td>1</td>
<td>568</td>
<td>14.4–3166</td>
</tr>
<tr>
<td>SSRIs</td>
<td>Non–NA</td>
<td>3497</td>
<td>432</td>
<td>5</td>
<td>1157</td>
<td>376–2701</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>1694</td>
<td>200</td>
<td>1</td>
<td>500</td>
<td>12.3–2786</td>
</tr>
<tr>
<td>TCAs</td>
<td>Non–NA</td>
<td>2347</td>
<td>245</td>
<td>3</td>
<td>1225</td>
<td>253–3579</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>740</td>
<td>71</td>
<td>0</td>
<td>0</td>
<td>0–5196</td>
</tr>
</tbody>
</table>

NA indicates North America.

*Reference courtesy of to Robert T. O’Neill, Ph.D., Senior Statistical Advisor to FDA
FDA’s review of RCT’s generally involves evaluation of study results (statistical analyses) according to region, and maybe country - often difficult to interpret*

- Evaluate the study data and the conduct and key metrics of quality
- Evaluate statistical displays of key sources of variation, bias and uncertainty
- Regional and site outcomes evaluated:
  - Dropouts, differences in response rates, outcomes, covariates, exposures, follow-up, concomitant drugs
- Possibly intrinsic factors (markers, gender, ethnicity) or possibly extrinsic factors, recruitment patterns, medical support system, standards of care (e.g., aspirin in T vs. C study)

*Slide Courtesy of Robert T. O’Neill, Ph.D., Senior Statistical Advisor to FDA
Interpretation of the global estimate and region specific estimates is challenging and the causes for heterogeneity are usually unknown.

Intrinsic or Extrinsic factors and their evaluation

*Slide courtesy of Robert T. O’Neill, Ph.D., Senior Statistical Advisor to FDA
Differences in treatment effects are expected.

Too much heterogeneity is problematic.

Issue - What to make of it?
Are these treatment differences real and are they systematic in the sense that treatment effects are consistently better or worse in the U.S. and what are the reasons for it?

*Slide courtesy of Robert T. O’Neill, Ph.D., Senior Statistical Advisor to FDA*
Region-Specific/Site-Specific Interim Monitoring of:

- Primary and secondary endpoint rates
- Demographic Characteristics
- Premature Withdrawal Rate
- Adverse event rates
- Adherence to Protocol
- Adherence to Treatment
- Data Timeliness
- Data Cleanliness
- Important Concomitant Medication Use
Case Study of DSMB Role in MRCT
Scenario

• Premature infants may be born with inadequate surfactant coating of the lungs
  – As compared to term infants, premature infants have higher incidence of:
    • Respiratory distress syndrome
    • Bronchopulmonary dysplasia
    • Mortality
• Exogenous surfactants exist which may be given to premature infants on birth (prophylactically) to reduce risk of above morbidities/mortality
Scenario

- Current surfactants are formulated from animal proteins (bovine, porcine)
  - They are effective in reducing morbidity/mortality rate
- A sponsor was interested in developing a new synthetic surfactant
  - Not expected to be more effective than current surfactants in reducing morbidity/mortality
  - Potential benefit: No animal microorganisms that may cause problems later ("mad cow")
Scenario

- Sponsor conducted a multi-regional non-inferiority clinical trial to compare the experimental (E) synthetic surfactant to a control animal-derived surfactant (C)

- Primary Efficacy Outcome:
  - Alive and without bronchopulmonary dysplasia (BPD) 28 days after birth
  - Trial’s non-inferiority objective: prove primary outcome rate of E is <13% lower (worse) than that of C (13% is the “non-inferiority margin”)
    - Sponsor did not try to prove E was “superior” to C because sponsor did not believe E was “superior” to C
Scenario

- Regions involved:
  - UK
  - Poland
  - Hungary
  - Spain
  - Portugal
  - France
  - USA
  - Canada
Scenario

• Study was designed assuming outcome rate of 55% in each treatment group
  – Assumption of 55% was based on a previously published study (1988) of the control group C
  – 305 subjects/group yielded 90% power to show non-inferiority of E to C when using a margin of 13%.

• DSMB periodically inspected data for safety issues
  – No decisions regarding study conduct (e.g., stopping for overwhelming efficacy, futility, increasing sample size) were to be made based on efficacy
  – However, DSMB was provided with efficacy outcome rates for both treatments combined.
Interim Results

• After 229 subjects were enrolled, an interim look at the primary efficacy outcome rate was performed.

• The primary outcome rate (alive without BPD) was $82/229 = 36\%$
  – Rate was similar in both groups

• Why was DSMB concerned?
  – Outcome rate was markedly lower than that anticipated in study design stage (55%)
Interim Analysis

• What were the reasons?
  – Sample simply not representative of the population?
  – Were some regions lowering the outcome rate?
  – Demographic characteristics?
  – Other?
**Interim Analysis**

- DSMB inspected outcome rates per region (not originally supplied to DSMB):

<table>
<thead>
<tr>
<th>Region</th>
<th>Primary Efficacy Outcome Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>8/15 (53%)</td>
</tr>
<tr>
<td>Poland</td>
<td>3/27 (11%)</td>
</tr>
<tr>
<td>Hungary</td>
<td>30/54 (56%)</td>
</tr>
<tr>
<td>Spain</td>
<td>10/44 (23%)</td>
</tr>
<tr>
<td>Portugal</td>
<td>4/8 (50%)</td>
</tr>
<tr>
<td>France</td>
<td>20/47 (42%)</td>
</tr>
<tr>
<td>USA</td>
<td>6/29 (21%)</td>
</tr>
<tr>
<td>Canada</td>
<td>1/5 (20%)</td>
</tr>
</tbody>
</table>

4 of 8 regions had outcome rates markedly lower than anticipated.
Interim Analysis

• Upon further inspection, it was noticed that average birth weight for the study was 939g.

• Why was this important?
  – Mean birth weight for the published study with the 55% outcome rate (on which the design of this study was based) was ~1200g.
  • Birth weight of premature infants is positively related to probability of surviving without BPD.
**Interim Analysis**

- Birth weights by region were inspected:

<table>
<thead>
<tr>
<th>Region</th>
<th>Primary Efficacy Outcome Rate</th>
<th>Median Birth Weight (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>8/15 (53%)</td>
<td>925</td>
</tr>
<tr>
<td>Poland</td>
<td>3/27 (11%)</td>
<td>961</td>
</tr>
<tr>
<td>Hungary</td>
<td>30/54 (56%)</td>
<td>975</td>
</tr>
<tr>
<td>Spain</td>
<td>10/44 (23%)</td>
<td>922</td>
</tr>
<tr>
<td>Portugal</td>
<td>4/8 (50%)</td>
<td>928</td>
</tr>
<tr>
<td>France</td>
<td>20/47 (42%)</td>
<td>909</td>
</tr>
<tr>
<td>USA</td>
<td>6/29 (21%)</td>
<td>928</td>
</tr>
<tr>
<td>Canada</td>
<td>1/5 (20%)</td>
<td>994</td>
</tr>
</tbody>
</table>
Interim Analysis

Outcome Rate (%) vs. Mean Birth Weight (g)
Interim Analysis

• Outcome rate vs. region-specific mean birth weight: No real relationship seen
  – However all mean birth weights < 1200g.
  – Several regions achieved ~55% outcome rate despite the low birth weight.

• It was recommended to sponsor to investigate this issue further because it was felt FDA would surely be concerned.
Epilogue

• FDA did indeed have a concern about low outcome rate.

• Sponsor prepared a white paper for FDA showing:

  – “Literature showed that the key events in preterm neonatal studies such as all-cause mortality and BPD or CLD (chronic lung disease) were highly correlated with the birth weight for VLBW neonates [1, 2, 3].”

• White paper also showed:
  
  - Birth weights in sponsor’s study ranges from 586g to 1560g.
  
  - While region-specific mean birth weights were not related to region-specific outcome rate, in general, within study, birth weights were related to outcome rate using logistic regression:
    
    • Odds of alive w/out BPD increased 77% per 100g increase in birth weight (95% CI of 49% to 112%, p<0.001).
    
    • Model fit the data very well
Epilogue

• Also, note that 1988 study on which 55% assumption was based was performed primarily in Western Europe
  - Netherlands, Ireland, Germany, Sweden, France, Italy
  - 3 of the 4 Western Europe countries in sponsor’s study had an outcome rate of 42%-56%.
• At the end of trial, non-inferiority of E to C was proven, but again with an overall event rate of approximately 35% in each treatment group

• Nevertheless, the FDA considered this a flawed trial and did not approve
  – (In addition to very low primary outcome rate, FDA felt that the non-inferiority margin was too large, especially if outcome rate is only 35%, and that the 1988 reference used to obtain the estimate of a 55% outcome rate was too “old” and that the true outcome rate should maybe now be larger)
• FDA recommendation: design a new study, using the lessons learned
  – Incorporate variability in outcome rate seen (a) across regions and (b) across patient weights when determining expected event rates, sample size
  – Easier said than done. It involves several assumptions:
    • Outcome rate in each region; though past data exists to make this possible, there is no guarantee this is the truth, especially for the smaller regions
    • Expected sample size in each region
    • Distribution of patient weights in study
  – If these assumptions are incorrect, results could be held in question
Conclusion

- MRCTs provide a way to conduct a clinical trial over widespread demographic
- MRCTs could show real or spurious treatment effect differences across regions, or real or spurious differences in outcome rates across regions
- Need to assess cause of differences if possible
- Though often not done, DSMBs should consider inspecting event rates by region
  - DSMB can have a role to help guide sponsor in detecting these differences and assessing cause of differences
DSMB DILEMMAS IN INTERNATIONAL TRIALS: A CASE STUDY

Susan S. Ellenberg, Ph.D.
Perelman School of Medicine at the University of Pennsylvania

Society for Clinical Trials Annual Meeting
Boston, MA
May 19, 2013
CASE STUDY

- Multinational trial, but within southern Africa
- “Multinational issues” less with managing the participating sites than with managing different perceptions and beliefs of DSMB members, study leadership and study sponsor
- Intensifying discussions
  - Conduct of trial in developing countries
  - Study population highly vulnerable (infants infected with HIV)
• National Institute of Allergy and Infectious Diseases (NIAID) sponsors many clinical trials related to HIV/AIDS worldwide
• From 2005-2010 I served on a DSMB for NIAID-sponsored trials of treatment for HIV conducted in southern Africa
• Lead investigators were in diverse locations
  — South Africa
  — US
  — Great Britain
• DSMB interacted with lead statistical and clinical investigators as well as NIAID staff
EARLY TRIAL

- DSMB was asked to monitor a study comparing treatment strategies for HIV-infected infants being treated in South Africa.

- Trial question: when to start antiretroviral therapy, and how long to treat.

- Primary endpoint: death or failure of therapy (based on CD4 levels, clinical outcomes or treatment toxicity).

- Lead investigators
  - Clinician: South African physician
  - Statistician: British (London)
DSMB COMPOSITION

- DSMB was geographically diverse, with representation from US, UK, South Africa, Botswana, Ghana, Zimbabwe
- Members included clinicians with experience treating HIV/AIDS, biostatisticians, bioethicists
- Not all members had prior DSMB experience
- Few had had prior involvement with each other, study sponsor or lead investigators
STANDARD OF CARE: SOUTH AFRICA

- National guidelines in Africa at that time was to delay treatment until clinical symptoms appeared or immune system became highly depressed.
- Modifying guideline to begin treatment earlier would have huge economic implications.
- Extremely strong findings would be required to persuade national authorities to revise guidelines because of drug costs.
In 2005 there was an emerging view that HIV-infected infants should receive treatment, regardless of symptoms or immunological/viral parameters—HIV infection known to progress more rapidly in infants.

The 2005 guidelines recommended treatment for all infected infants with symptoms or depressed CD4 counts, and recommended that treatment be considered for all others.

Guidelines noted that some experts recommended universal treatment.
STUDY DESIGN

- Three arm design
  - Defer treatment until child qualified according to national guidelines
    - CD4% < 20 (changed to 25 partway through trial), or
    - Development of clinical symptoms
  - Begin treatment immediately following diagnosis and continue for 40 weeks
  - Begin treatment immediately following diagnosis and continue for 96 weeks
EARLY STOPPING PLAN

- The study protocol specified a stringent Haybittle-Peto early stopping plan to the DSMB
  - Require a treatment effect significant at the 0.001 level at any interim analysis
  - Presentation of pooled mortality data only at reviews when formal interim analysis not scheduled

- DSMB was not fully comfortable with this plan
  - Acknowledged need for more data to have impact on national treatment policies
  - Shared emerging concerns about leaving HIV-infants untreated
POSSIBLE ALTERNATIVE PLAN

- In US, most traditional approach to early stopping is the O’Brien-Fleming design
  - Difficult to stop very early, but easier as more data accumulate
  - At any point after the first or second look, would be more permissive of early stopping
- Study team argued that the more conservative stopping plan was essential
  - Very important to “get it right”
  - Strongest possible evidence would be needed to support changing guidelines
Fig. 15-4 Three group sequential stopping boundaries for the standardized normal statistic \(Z_i\) for up to five sequential groups with two-sided significance level of 0.05.
# HAYBITTLE-PETO VS O'BRIEN-FLEMING

## 3 Tests

<table>
<thead>
<tr>
<th></th>
<th>H-P</th>
<th></th>
<th>O-F</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H-P</td>
<td>0.001</td>
<td></td>
<td>0.001</td>
<td>0.05</td>
</tr>
<tr>
<td>O-F</td>
<td>0.0006</td>
<td>0.0151</td>
<td>0.0472</td>
<td></td>
</tr>
</tbody>
</table>

## 5 Tests

<table>
<thead>
<tr>
<th></th>
<th>H-P</th>
<th></th>
<th>O-F</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H-P</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>O-F</td>
<td>5x10^{-6}</td>
<td>0.0013</td>
<td>0.0085</td>
<td>0.0228</td>
</tr>
</tbody>
</table>
Fig. 15-4 Three group sequential stopping boundaries for the standardized normal statistic ($Z_i$) for up to five sequential groups with two-sided significance level of 0.05.
DIFFERENT VIEWS ON EARLY STOPPING

- Cultural differences in approach to monitoring trials with the possibility of stopping early for efficacy
  - Perspective 1: trial should not stop early unless conclusions are very unlikely to change
  - Perspective 2: trial should not stop early unless data are sufficiently strong that they will lead to change in medical practice

- Researchers in US tend to take Perspective 1: researchers in UK tend to take Perspective 2
TRIAL INITIATED

- DSMB did not insist on any change in proposed design
  - Accepted Haybittle-Peto monitoring plan (with some reluctance), recognizing that if results showed need for immediate treatment, they would have to be very strong to lead to change in practice

- Within first year of study, and several months prior to first planned analysis, NIAID program staff asked for “emergency” DSMB meeting
NIH CONCERNS

- Program staff (at that time) received reports of deaths in real time, unblinded to treatment arm
- Shared view of some on DSMB that early treatment was likely effective; worried about infants on untreated arm
- Early data showed 3 deaths on untreated arm, no deaths on either arm receiving treatment
DSMB WAS UNANIMOUS

- Although there had been mixed views about the Haybittle-Peto stopping rule, no one on the DSMB thought these data warranted action
  - With 3 deaths, in a 3-arm trial, if there were no treatment effect would expect 1 death in each arm
  - 3-0-0 is not different enough from 1-1-1 to cause panic—could easily be due to random fluctuation
Although there had been mixed views about the Haybittle-Peto stopping rule, no one on the DSMB thought these data warranted action

- With 3 deaths, in a 3-arm trial, if there were no treatment effect would expect 1 death in each arm
- 3-0-0 is not different enough from 1-1-1 to cause panic—could easily be due to random fluctuation

But...the 3 deaths were all in the delayed treatment arm, fueling concerns about possible early stopping in the future
The Board was now even more uncomfortable about the monitoring approach, given the suggestion (however slight) of an emerging difference.

- Asked that a subset of the Board review comparative data every 3 months, with the possibility of referral to full Board
MORE ON THE STOPPING RULE

- Again raised the issue of the stringency of the early stopping rule
  - Asked for confirmation that the study team remained comfortable with that rule
  - Asked the study team to propose criteria for unscheduled interim analysis; next scheduled review not for another year
Third full board review; subset of board had been reviewing quarterly mortality data

At this meeting the 0.001 level was not yet reached for the protocol-specified primary outcomes: mortality/treatment failure in each immediate treatment arm compared to deferred treatment arm
  - Deferred vs 40 weeks: p=0.008
  - Deferred vs 92 weeks: p=0.01

Comparison of the untreated arm with the two immediate treatment arms combined yielded p<0.001
RECOMMENDATION

- Pre-specified stopping boundary not crossed
- Nevertheless, considering these data and other data available at the time, DSMB was not comfortable continuing study as designed
- DSMB recommendation
  - Terminate untreated arm
  - Refer all infants in that arm to clinic for re-evaluation and consideration of treatment
  - Continue study of 2 immediate treatment arms
- Recommendation accepted by study team and by NIAID
**RESULTS**

*NEJM* 359:2233-44, 2008

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Early ARV (N=252)</th>
<th>Deferred ARV (N=125)</th>
<th>Total (N=377)</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death</td>
<td>10 (4%)</td>
<td>20 (16%)</td>
<td>30 (8%)</td>
<td>0.24 (0.11, 0.51)</td>
</tr>
<tr>
<td>Non-fatal treatment failure</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Total endpoints</td>
<td>11</td>
<td>21</td>
<td>32</td>
<td>0.25 (0.12, 0.51)</td>
</tr>
</tbody>
</table>
CONSEQUENCES

- Within a year of recommendation, WHO guidelines were changed.
- Strong recommendation for initiating antiretroviral treatment for all HIV-infected children under 12 months of age regardless of clinical symptoms or immune status.
There will inevitably be tensions when monitoring studies in vulnerable populations in which mortality is an outcome.

These tensions can be exacerbated when there are differing views among DSMB members, investigators and sponsors as to the level of evidence required to stop a study early:

- Consideration of safety of children in study
- Consideration of safety of future children whose treatment may depend on study outcome
- Consideration of evidence required to influence policy/funding decisions
In trials when participants are from different settings, it is critical for DSMB members to understand and appreciate:

- National treatment policies in the areas in which the trial is being conducted
- Relevant regulatory policies in the areas in which the trial is being conducted
Although there can be differing views among members of any DSMB, differences may be more likely when trial sponsor, trial investigators and DSMB members start off with differing perspectives on what should lead to stopping a study early.

These differing perspectives may be associated with different geographical settings.

Critical for DSMB, sponsor and investigators to be in agreement on monitoring boundaries, criteria for early stopping.
Considerations for DMC Decision-Making: Beyond Stopping Boundaries

Discussion of ASCOT-BPLA: Anglo-Scandinavian Cardiac Outcomes Trial–Blood Pressure Lowering Arm

Charles Knirsch MD, MPH
Head, Clinical Research,
Pfizer Specialty Care
Points Covered in this Talk

• Perspectives for a DMC to consider in stopping a trial

• Primary, Secondary and Accumulated Data (previous trials)

• Interactions between DMC, Executive Committee and Sponsor
ASCOT Rationale

- CHD incidence remains a major unresolved problem in BP management
- High prevalence of dyslipidaemia in hypertensive patients
- Combinations of risk factors synergistic for CHD
- No trial has specifically addressed benefits of lipid lowering in primary prevention of CHD in hypertensive patients not conventionally deemed dyslipidaemic
ASCOT Study Design

18,000 patients

R

9000 β-blocker ± diuretic

9000 CCB ± ACEI

500 TC ≤ 6.5 mmol/L (≤ 250 mg/dL)

4000 TC > 6.5 mmol/L (> 250 mg/dL)

4000 TC > 6.5 mmol/L (> 250 mg/dL)

5000 TC ≤ 6.5 mmol/L (≤ 250 mg/dL)

R

500 open lipid lowering

4500

4500

500 open lipid lowering

R

2250 statin

2250 placebo

2250 statin

2250 placebo

8000 open lipid lowering

ascot

Lipid Lowering Arm of Study Stopped

• The DSMB in September 2002 reported that in the lipid arm of ASCOT there was a highly significant reduction in the primary end point as well as a significant reduction in stroke.

• The DSMB recommended that the double-blind cholesterol lowering study treatment arm be terminated since the results crossed the pre-defined Peto stopping boundaries of the trial.

• The Executive Committee endorsed the recommendation of the DSMB, and the lipid arm was closed after a median follow-up period of 3.3 years.
ASCOT: Anglo-Scandinavian Cardiac Outcomes Trial

ASCOT — multicenter, international trial comparing treatment regimens

Study 1: ASCOT-LLA

- Stopped after 3.3 years for benefit

Study 2: ASCOT-BPLA

- Prospective, randomized, open, blinded endpoint (PROBE) design comparing two antihypertensive regimens in the total ASCOT patient population

The following slides focus on the ASCOT-BPLA arm

ASCOT-BPLA: Study design

**Design:** Double-blind, randomized, blinded adjudication

**Population:** N = 19,257 with hypertension and ≥3 other CV risk factors

**Treatment:**
- Amlodipine 5–10 mg ± perindopril 4–8 mg prn (n = 9639)
- Atenolol 50–100 mg ± bendroflumethiazide 1.25–2.5 mg/potassium prn (n = 9618)

**Primary outcome:** Nonfatal MI (including silent MI) and fatal CHD

**Secondary outcome:** All-cause mortality, stroke, nonfatal MI (excluding silent MI), all coronary events, CV events/procedures, CV mortality, fatal/nonfatal HF
ASCOT-BPLA: Treatment algorithm for BP targets

BP medication titrated to achieve target:
No diabetes: <140/90 mm Hg
Diabetes: <130/80 mm Hg

19,342 patients
40–79 y
with
UNTREATED
SBP ≥160 mmHg
and/or
DBP ≥100 mmHg
OR
TREATED
SBP ≥140 mmHg
and/or
DBP ≥90 mmHg

In each arm, pts with total cholesterol ≤6.5 mmol/L randomized to atorvastatin (10 mg) or placebo daily (n = 10,297)

Amlo = amlodipine; Peri = perindopril; Doxa = doxazosin GITS (Gastrointestinal Transport System); BFZ = bendroflumethiazide

Data for Interim Recommendation #1
November 2003

• Secondary endpoint: Trend of fewer strokes in amlodipine arm (230 vs 339, P= 0.00004)

• Primary endpoint of non-fatal MI and fatal CHD had fewer events in amlodipine arm (313 vs 354, P=0.14)

• Consistent with literature that anti-hypertensive therapy reduced stroke more readily than coronary heart disease

  – Pocock SJ, Clinical Trials 2006; 3:513-521
DMC Recommendation and Actions

• DMC recommended stopping the trial

• Trial Executive Committee consulted external experts to consider stopping the trial on the secondary endpoint of stroke
  – Noted un-adjudicated events

• Trial continued for 11 additional months with an emphasis on swift adjudication of events and monthly reporting to the DMC
  » Pocock SJ, Clinical Trials 2006; 3:513-521
October 2004: ASCOT-BPLA: Reduction in primary outcome (nonfatal MI and fatal CHD)

HR = 0.90 (95% CI, 0.79–1.02)
RRR = 10%
P = 0.1052

Number at risk
Amlodipine-based regimen (429 events) 9639 9475 9337 9168 8966 7863
Atenolol-based regimen (474 events) 9618 9470 9290 9083 8858 7743

October 2004: ASCOT-BPLA: Reduction in fatal and nonfatal stroke

Number at risk
Amlodipine-based regimen
(327 events)
9639 9483 9331 9156 8972 7863
Atenolol-based regimen
(422 events)
9618 9461 9274 9059 8843 7720

Proportion of events (%)

Time (years)

Atenolol-based regimen*
Amlodipidine-based regimen†

HR = 0.77 (95% CI, 0.66–0.89)
RRR = 23%
P = 0.0003

ASCOT-BPLA: Reductions in BP over time

Mean difference = 1.9, P < 0.0001

Mean difference = 2.7, P < 0.0001

ASCOT-BPLA: Additional reductions in group receiving the amlodipine-based regimen

<table>
<thead>
<tr>
<th>Secondary endpoints</th>
<th>Amlodipine-based* (n = 9639)</th>
<th>Atenolol-based † (n = 9618)</th>
<th>Rate/1000 patient years</th>
<th>Rate/1000 patient years</th>
<th>Amlodipine-based better</th>
<th>Atenolol-based better</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonfatal MI (excluding silent) + fatal CHD</td>
<td>7.4</td>
<td>8.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total coronary endpoint</td>
<td>14.6</td>
<td>16.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total CV events and procedures</td>
<td>27.4</td>
<td>32.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>13.9</td>
<td>15.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>CV mortality</td>
<td>4.9</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Fatal/nonfatal stroke</td>
<td>6.2</td>
<td>8.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fatal/nonfatal HF</td>
<td>2.5</td>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tertiary endpoints</th>
<th>Amlodipine-based* (n = 9639)</th>
<th>Atenolol-based † (n = 9618)</th>
<th>Rate/1000 patient years</th>
<th>Rate/1000 patient years</th>
<th>Amlodipine-based better</th>
<th>Atenolol-based better</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of diabetes</td>
<td>11.0</td>
<td>15.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Development of renal impairment</td>
<td>7.7</td>
<td>9.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

Unadjusted risk reduction

ASCOT-BPLA: Overall results

• Study stopped prematurely after 5.5-year median follow-up because of higher death rate in assigned atenolol-based-regimen group

• Group receiving amlodipine-based regimen had non-significant 10% reduction in primary outcome (nonfatal MI plus fatal CHD) and significant reductions in nearly all secondary CV endpoints and new-onset diabetes

Some Issues from the DMC Perspective

• The nature of the risk or benefit that the DMC must consider
  • stroke as a secondary endpoint and the broader primary endpoint of MI and fatal CHD

• Consequences to research participants and to future patients if no further information is obtained from the randomized controlled trial
  – If stopped for stroke, is the relationship between BP lowering and CHD established with these therapeutics?

• Ethical issue of the weight of the information considered by the DMC and the benefit-risk disclosed in the consent and consequent reasonable understanding of research participants
DMC Considerations

• The safety and interests of the participants

• The obligation to share important new information with participants and researchers

• The scientific validity of potentially important new information
  • Guidelines and changes to standard care

• The effect that stopping the study will have on the ability to answer important scientific questions affecting future patients
Considerations for Assessing New Data from ASCOT

• Likelihood that benefits to participants (stroke prevention) are related to the experimental intervention and BP changes
  – Regional differences in perception of weight of evidence and ability to change practice

• Sponsor – Regulator Perspective and Agreements / Interactions
Summary

• DMC experience, autonomy and recommendations critical
• Executive committee interactions with ability to consult external experts
  – Regulators?

• Learn from Case Studies, Discuss and Build Capacity
  – Clinical Trials Transformation Initiative
    • (https://www.ctti-clinicaltrials.org/)
  – Multi-regional Clinical Trials Initiative
    • (http://mrct.globalhealth.harvard.edu/)
  – Academic work
Multiregional Clinical Trials: How should DMCs look at interim data?

Janet Wittes
Statistics Collaborative
Outline

• Credo (aka, my philosophy)
• Why do we study drugs cross nationally?
• Is region (country) just another subgroup?
• If not, why not?
• How should DMCs look at multinational data?
Why study disease in many countries?

• Good reasons
  ▫ We believe in homogeneity of effect
  ▫ We are interested in population heterogeneity

• Bad reasons
  ▫ Desire to carry out trials where it is cheap
    • Recruitment easy
    • People will adhere
    • Costs low
  ▫ Desire to study people without other meds
But we know to worry about subgroups

![Figure 2: Scatterplot of the Relative Risk for Fatal/Nonfatal Cardiovascular Disease Events By Subgroup Size (presented on a log scale)](image)

- Overall Ln(Relative Risk) for the entire SHEP Trial = -0.37 (indicated by the ----), corresponding to a Relative Risk of 0.69
Why are regional subgroups different from all other subgroups?

- Even a given country is not homogeneous
  - (in the US, e pluribus unum)
- And study populations are rarely representative
Disease factors that differ by country

- Genetic diseases: different genotypes
- Infectious disease: different organisms & seasons
  - Follow flu from China to Southern to Northern Hemi
  - Study malaria where it is (Willie Sutton rule)
  - Select different degrees of endemicity
- Chronic disease: different stage of disease
  - cervical cancer: Pap smear countries vs. other
  - heart failure US+W Europe vs Russia+E Eur
Belimumab for lupus

Response Rate
(SELENA-SLEDAI improvement 4 or more points, no clinically significant worsening in BILAG or Physician’s Global)

Placebo (N=275) 34%
Hi (N=273) 43%

FDA briefing document 19-Oct-10
Belimumab for lupus

<table>
<thead>
<tr>
<th>Region</th>
<th>Placebo (N=275)</th>
<th>Hi dose (N=273)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>34%</td>
<td>43%</td>
</tr>
<tr>
<td>W Eur/Isr</td>
<td>23%</td>
<td>51%</td>
</tr>
<tr>
<td>E. Eur</td>
<td>42%</td>
<td>53%</td>
</tr>
<tr>
<td>LA/SA</td>
<td>57%</td>
<td>53%</td>
</tr>
</tbody>
</table>

*FDA briefing document 19-Oct-10*
Belimumab for lupus

<table>
<thead>
<tr>
<th>Region</th>
<th>Placebo (N=275)</th>
<th>Hi dose (N=273)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>34%</td>
<td>43%</td>
</tr>
<tr>
<td>W Eur/Isr</td>
<td>23%</td>
<td>51%</td>
</tr>
<tr>
<td>E. Eur</td>
<td>42%</td>
<td>53%</td>
</tr>
<tr>
<td>LA/SA</td>
<td>57%</td>
<td>53%</td>
</tr>
<tr>
<td>USA/Can</td>
<td>32%</td>
<td>35%</td>
</tr>
</tbody>
</table>

*FDA briefing document 19-Oct-10*
Approved in the US and Europe

“It’s unsettling to me to contemplate approval of a drug for patients in the United States, for practitioners in the United States, that does not seem to work for patients in the United States,” said Dr. David Blumenthal, a committee member and rheumatologist at Case Western Reserve University. He voted against approval.
Rephrase: Why is this subgroup (region) different from all other subgroups (or is it)?

- If it’s not, we should (nearly) ignore different effects
- It feels different (different results by country)
  - Disease factors
  - Population factors
  - Quality of care
- And what should a DMC know or do?
A trial sponsored by CVD-Mali

- Centre pour les Vaccins en Développement
- 2001 concordat between
  - Ministry of Health, Mali
    - Dr. Fatoumata Nafo-Traore, Minister of Health
  - U Maryland, Baltimore
    - Prof Myron M. Levine, Director CVD-Maryland
Mission of CVD-Mali

- Prevent, control, & treat endemic & epidemic diseases in Mali
- Especially those currently or potentially vaccine-preventable
- In so doing --
  - to train cadres of Malian specialists who will expand the work in the future
Background of the trial

- CDC: pregnant women should get inactivated flu vaccine
- Developing countries: no routine flu vaccination
- 2008 Bangladesh – flu vaccine in pregnancy reduced
  - Influenza illness by 63% in infants <6 months of age
  - All febrile respiratory illness by 1/3 in mothers and infants
- 2010 CVD-Mali: surveillance of pregnant women & their infants
Randomized trial (n~4000)

- Arms
  - Influenza vaccine (TIV Inactivated)
  - Quadrivalent meningococcal conjugate
- Primary objectives
  - Incidence of lab-confirmed influenza among infants up to 6 mo
    - In ITT pop
    - With mothers immunized ≥14 days before delivery
DSMB: safety of mothers and infants

- Where are we from?
  - US -3
  - UK -1
  - Ghana -1
  - Mali -1
  - South Africa -1
What do we need to know...

- The usual suspects
  - The disease
  - The study design
  - Our role
  - ....
But also: where is Mali?
Some salient facts about Mali?

• Population ~12.5 million
• Bamako, the capital, around 1.1 million
• Landlocked, resource poor
• Economy: One of the 10 poorest countries
• Seasonal malaria
• HIV prevalence ~2%
Field trip

• The DSMB visited Mali
  ▫ the clinic sites
  ▫ the university hospital
• Visit was sobering
  ▫ We realized we had to be frugal in our questions
  ▫ We left with a sense of awe about CVD-Mali
• Plane ride home was delayed – no fuel in Mali
Our reviews

• Have been reviewing on a regular basis
  ▫ Recruitment
  ▫ Major adverse events in mothers and infants
  ▫ Rates of prematurity and stillbirth
The fact that this large trial (> 4200 pregnant women enrolled as of this week) could have reached this milestone and begin the countdown to completion is an extraordinary feat considering all the political upheaval (coup d’état, counter coup, curfews, urban fighting), periodic strikes and protests that interfered with vehicular transport, invasion of Northern Mali by mercenaries returning from Libya and infiltration by hundreds of Al Qaeda-affiliated jihadists who for several months took over the towns of the North, terrorizing the local populations.

Next came the expedition into the North by French troops and liberation of the major towns of Gao, Kidal and Timbuktu. That surveillance and follow-up of enrolled subjects and further enrollment could have continued almost unabated is a tribute to the professionalism, dedication, competence, leadership, commitment, bravery and resourcefulness of Samba Sow and Mili Tapia and their colleagues. Incredibly, there were only ~ 3 weeks during the past 13 months when CVD-Mali field teams were unable to maintain follow-up because of curfews and urban fighting.

Those of you who visited CVD-Mali on the occasion of the first DSMB meeting will recall that the home base of the parachute regiment (who were loyal to deposed President Amadou Toumani Touré) sits in part adjacent to and in part directly across the road from CVD-Mali headquarters. During the time of the counter-coup there was vicious fighting between the parachute regiment loyal to the democratically-elected deposed President and the regiment of the perpetrators of the coup.
DSMB should understand population factors

- Diet
- Risk factors
  - Smoking
  - Drinking
  - Comorbidities
- Racial (genetic) and ethnic (cultural) differences
- Potential political problems
Treatment factors

• Medical care in the countries involved
• Standard of care
  ▫ Time of diagnosis
  ▫ Use of drugs
  ▫ Surgical interventions
• Adherence to protocol
How does this relate to interim monitoring?

- Do regional analyses
- Do other subgroup analyses
- If you stop early, it is what it is
The PROTECT Study

- Rolofylline + placebo both + loop diuretic
- Heart failure signs and symptoms
Design (hospitalized heart failure)

- 600 patients (later 2000)
- 2:1 active to placebo
- 75 sites (US, Israel, E and W Europe, Russia)
- 3 day infusion of drug or placebo
- Outcome is day 2 and 3
  - Other measures at Day 7, 60, and 180
  - Study ends when last patient has 60 days f-up
Primary outcome: 3 category variable

- **Success**
  - Dyspnea Day 2 & 3 moderately or markedly better
  - Not a treatment failure
- **No change – not a success or a failure**
- **Failure**
  - Worsening symptoms
  - Death, hospital readmission, or other bad things
Safety concern

- Drug is an adenosine A1 receptor agonist
- Known to lower seizure threshold
The PROTECT Study

- Rolofylline + placebo both + loop diuretic
- Heart failure signs and symptoms
Design (hospitalized heart failure)

- 600 patients (later 2000)
- 2:1 active to placebo
- 75 sites (US, Israel, E and W Europe, Russia)
- 3 day infusion of drug or placebo
- Outcome was at day 2 and 3
  - Other measures at Day 7, 60, and 180
  - Study ends when last patient has 60 days f-up
Primary outcome: 3 category variable

- **Success**
  - Dyspnea Day 2 & 3 moderately or markedly better
  - Not a treatment failure
- **No change – not a success or a failure**
- **Failure**
  - Worsening symptoms
  - Death, hospital readmission, or other bad things
Safety concern

• Drug is an adenosine A1 receptor agonist
• Known to lower seizure threshold
Results from Russia differed from rest of trial

- The DMC looked carefully at the data to make sure that
  - The patients were safe
  - The data would be interpretable at the end
MERIT-HF

• Randomized, double-blind, placebo-controlled trial
  • Symptomatic heart failure
  • Metoprolol (different doses depending on NYHA class)
  • ~2000 participants / group
  • 13 European countries + US

• Co-primary outcomes (time to) - either
  • total mortality OR
  • combined endpoint of total mortality or all-cause hospitalizations

Study stopped at 2\textsuperscript{nd} interim analysis

• 50% information; \( p < 0.001 \)
• Mean follow-up: 1 year
• Deaths
  ▪ Metoprolol: 145
  ▪ Placebo: 217
  ▪ Relative risk: 0.66
  ▪ 95% CL: (0.53, 0.81)
## MERIT-HF

<table>
<thead>
<tr>
<th>Region</th>
<th>Relative Risk</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.66</td>
<td>(0.53, 0.81)</td>
</tr>
<tr>
<td>USA</td>
<td>1.05</td>
<td>(0.71, 1.56)</td>
</tr>
<tr>
<td>Ex-US</td>
<td>0.55</td>
<td>(0.43, 0.70)</td>
</tr>
</tbody>
</table>

Interaction p-value: 0.003
If the mortality endpoint is the most important among all endpoints, the US sub-population should be the most important subgroup in a multinational trial because the goal of the NDA submission is to gain approval for marketing the drug in the US. The efficacy outcome in this population must be examined carefully as part of the evaluation of the totality of the evidence and possible extrapolation of the efficacy evidence from foreign population[s] to [the] US population.
Conclusion: DMCs should routinely look at region

- But: look should be careful
  - Are the patients different in meaningful ways?
- Does the regional difference look like:
  - Chance?
  - Confounding?
- The DMC should ask for more analyses (or let the study continue) if the data are confusing