Supported by:

American Association for Cancer Research
American Society of Clinical Oncology
Susan G. Komen
GLOBAL HARMONIZATION OF DRUG DEVELOPMENT
Richard Pazdur, MD; FDA OCE
Prudence Scott, MD; Medex Consulting
Francesco Pignatti, MD; EMA
FRIENDS ANNUAL MEETING

PANEL THREE:
OPTIMIZATION OF EXPLORATORY RANDOMIZED TRIALS
FRIENDS ANNUAL MEETING

LISA LAVANGE, PHD
FDA
Panelists

- Lisa LaVange, PhD; FDA
- Eric Kowack, MS, MBA; Ignyta
- Amy McKee, MD; FDA
- Cyrus Mehta, PhD; Cytel
- Richard Simon, D.Sc; NCI
- Rajeshwari Sridhara, PhD; FDA
- Yanping Wang, PhD; Eli Lilly & Co.
Optimization of Exploratory Randomized Trials

Lisa M. LaVange, PhD
Director, Office of Biostatistics
CDER, FDA

Friends of Cancer Research Annual Meeting
November 16, 2016
Background

• Recent oncology research → new, highly effective therapies
  – May show large treatment effects early in development
  – Need to speed these new therapies to market while also ensuring sufficient rigor to allow regulatory scrutiny

• Specific challenges include:
  – Lack of planning in certain aspects of the trial can make results difficult to interpret, e.g., multiple looks at the data, multiple endpoints
  – Early evidence of effect may be seen in secondary endpoints, e.g., overall survival
  – Results can negatively impact ability of the next, larger pivotal trial to enroll
Session Objectives

• Better understand the problem through some recent examples
  – Focus is on early randomized trials designed to inform sponsors on a go/no-go decision
  – Options in reaction to exceptional results include trial expansion to determine if benefit is maintained or submission of data while possibly initiating pivotal trial

• Discuss options for FDA in interpreting early trial results
  – Are there statistical methods that can be applied post hoc to alleviate problems incurred with looser/unknown operating characteristics of early trials?

• Discuss options for sponsors to better plan for the possibility of exciting results in early trials
  – With care to not impede research by placing too many requirements on early trials
FRIENDS ANNUAL MEETING

Amy McKee, MD
FDA
Go/No-go Trials Supporting Approval: Case Studies

Amy McKee, M.D.
Deputy Office Director (acting)
Office of Hematology and Oncology Products
OND/CDER/FDA
Issue

• Small, activity-estimating, **randomized** trial reads out with unexpected results, perhaps in secondary endpoint

• Statistical plan written with less restrictive operating characteristics typical of phase 2 trial

• Regulatory conundrum: what to do with a marketing application based on this type of trial?
Case study #1: Olaratumab (Lartruvo™)

- Human Anti-PDGFRα Monoclonal Antibody
- Small, randomized phase 1b/2 trial (JGDDG)
  - Phase 2 portion randomized as add-on versus standard therapy in first-line setting
  - 133 patients with advanced STS not amenable to curative surgery/radiation
  - Primary endpoint: PFS (2-sided alpha=0.2)
    - Secondary endpoint: OS
**JGDG Trial Design**

- **Stratification factors:**
  - PDGFR-α expression
  - Prior lines of treatment
  - Histology
  - ECOG PS

- **130 Patients with Advanced STS**
  - Randomize 1:1

- **1° endpoint:** PFS

- **Olaratumab 15 mg/kg day 1 & 8**
  - Doxorubucin 75 mg/m² day 1

  - Treat x 8 21-day cycles

  - Olaratumab 15 mg/kg day 1 & 8

- **Disease evaluations:**
  - every 6 weeks x 4
  - then every 12 weeks
Case study #1: Olaratumab (Lartruvo™)
## Case study #1: Olaratumab (Lartruvo™)

<table>
<thead>
<tr>
<th></th>
<th>Olaratumab + Doxorubicin N=66</th>
<th>Doxorubicin N=67</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PFS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (months)</td>
<td>6.6 months [95% CI: 4.1, 8.3]</td>
<td>4.1 months [95% CI: 2.8, 5.4]</td>
<td></td>
</tr>
<tr>
<td>Hazard Ratio</td>
<td>0.67</td>
<td></td>
<td>0.06*</td>
</tr>
<tr>
<td><strong>OS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (months)</td>
<td>26.5 months [95% CI: 20.9, 31.7]</td>
<td>14.7 months [95% CI: 9.2, 17.1]</td>
<td></td>
</tr>
<tr>
<td>Hazard Ratio</td>
<td>0.46</td>
<td></td>
<td>0.0004</td>
</tr>
</tbody>
</table>

*2-sided alpha=0.1999*
Case study #2: Lenvatinib (Lenvima®)

• Small molecule tyrosine kinase inhibitor against VEGFR1, VEGFR2 and VEGFR3

• Small, randomized phase 1b/2 trial
  – Phase 2 portion randomized 1:1:1 to lenvatinib + everolimus, everolimus alone, or lenvatinib alone
  – 153 patients with renal cell carcinoma in second-line setting
  – Primary endpoint: PFS
    • Secondary endpoint: OS
Trial Design

Stratification factors:
- Hemoglobin
- Corrected serum calcium

150 Patients with Advanced STS

Randomize 1:1:1

- Lenvatinib 24 mg daily + Everolimus 5 mg daily
- Everolimus 10 mg daily
- Lenvatinib 24 mg daily

Disease evaluations:
- every 8 weeks

Treat to PD, unacceptable toxicity, consent withdrawal

1° endpoint: PFS
Case study #2: Lenvatinib (Lenvima®)
## Case study #2: Lenvatinib (Lenvima®)

<table>
<thead>
<tr>
<th></th>
<th>Lenvatinib + Everolimus N=51</th>
<th>Everolimus alone N=50</th>
<th>Lenvatinib alone N=52</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PFS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>14.6 months</td>
<td>5.5 months</td>
<td>7.4 months</td>
</tr>
<tr>
<td>Hazard Ratio vs evero alone</td>
<td>0.37* [95% CI: 0.22, 0.62]</td>
<td>0.57* [95% CI: 0.36, 0.91]</td>
<td></td>
</tr>
<tr>
<td><strong>OS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>25.5 months</td>
<td>15.4 months</td>
<td>19.1 months</td>
</tr>
<tr>
<td>Hazard Ratio vs evero alone</td>
<td>0.67* [95% CI: 0.42, 1.08]</td>
<td>0.80* [95% CI: 0.50, 1.27]</td>
<td></td>
</tr>
</tbody>
</table>

*p-values uninterpretable due to lack of pre-specified multiplicity adjustment
Conclusions

- Both olaratumab and lenvatinib received approval
- How can we prospectively design Phase 2 trials potentially to be both go/no-go and registration trials?
Bayesian analysis of unexpected survival “significance” in a randomized phase II trial

Richard Simon, D.Sc.
Many phase II trials are randomized because:
- New drug is combined with standard regimen
- Evaluation of progression free survival requires control group
- Toxicity better assessed with control group
## Differences Between Phase III and Randomized Phase II

<table>
<thead>
<tr>
<th></th>
<th>Phase II</th>
<th>Phase III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary endpoint</td>
<td>Progression-free survival</td>
<td>Survival</td>
</tr>
<tr>
<td></td>
<td>Response rate</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>Small - moderate</td>
<td>Large</td>
</tr>
<tr>
<td>Significance threshold</td>
<td>0.10 1-sided</td>
<td>0.05 2-sided</td>
</tr>
<tr>
<td>Analysis</td>
<td>Not-blinded</td>
<td>Blinded</td>
</tr>
<tr>
<td>Analysis plan</td>
<td>Flexible</td>
<td>Pre-specified</td>
</tr>
</tbody>
</table>
Surprises in RCT Outcomes

- Treatment difference for secondary endpoint
  - When OS is the secondary endpoint
- Treatment difference for subset of patients when there is no overall difference
- Treatment difference for a subset of clinical centers
- Un-expected serious toxicity from a list of examined toxicities
Some Problems With p Values for Data-driven Hypotheses

- The interpretation of p value as false-positive error requires pre-specification of hypothesis.
- Proper calculation of p value depends on sampling plan which may have been flexible or data driven; not fixed sample size as assumed.
- Same p value calculation for primary as for 10 secondary hypotheses. Only way of penalizing is by multiplicity correction.
Some Problems With p Values for Data-driven Hypotheses

- Based on null hypothesis of no treatment effect; does not reflect clinical significance
Why a Bayesian Model?

- Depends on data at analysis, not on data at earlier analyses
- Provides direct method for taking account of the fact that the outcome was for a secondary endpoint
- Can take into account clinical significance
Model

\[ \delta = \log \text{hazard ratio for survival of patient receiving test treatment } T \text{ relative to survival of patient receiving control } C \]

\[ \text{Hazard} \ (t) = \text{risk of death in interval} \ (t, \ t + \varepsilon) \text{ given that the patient is alive at time } t \]

Proportional hazard models assume that the hazard ratio is the same at all times \( t \).

\[ \text{hazard ratio} = \frac{Rx \ T \ \text{hazard}(t)}{Rx \ C \ \text{hazard}(t)} \]
Model

- hazard ratio = 1 means no treatment effect
- hazard ratio = 0.75 or 0.70 means there is a minimally clinically significant treatment effect
- hazard ratio ≤ 0.75 or ≤ 0.70 means that there is a treatment effect that is at least minimally clinically significant
- The true hazard ratio is unknown. The clinical trial provides us some information about it.
Model

- Using the data from the clinical trial, we would like to compute the probability that hazard ratio \( \leq 0.75 \) or 0.70

- Since it uses the data from the clinical trial, it is called a posterior probability

- In order to compute the posterior probability, we have to specify our prior probability about the hazard ratio
Our prior belief about the hazard ratio is not a number; it is a specification of possible hazard ratio values and strengths of belief we hold for those numbers.

People with very strong prior beliefs that hazard ratio=1 will need very strong evidence to change their beliefs.
If the clinical trial were a phase III trial being conducted because previous phase II trials indicated that the test treatment was effective, our prior belief that the hazard ratio = 1 might be about 50% or less.

For a phase II trial being planned to evaluate whether the treatment has an effect on PFS, a “surprising result” on OS means we have a greater prior belief that the hazard ratio for OS = 1.

For our calculations, we take the prior probability that the hazard ratio = 1 as 0.90
With this Bayesian model, we have placed the rest of the prior probability about the hazard ratio over the entire region < 1 using the negative part of a Gaussian distribution $N(0, \tau)$ for $\log$-hazard ratio.

We have found that the results are robust with regard to the specification of the standard deviation $\tau$. 
To Compute the Posterior Probability Distribution of the hazard ratio

- Specify the estimated hazard ratio based on the clinical trial data
- Specify the number of total deaths observed in the final analysis of the clinical trial
  - The sample variance of the observed log hazard ratio equals $4/\text{total deaths}$
- Specify your prior beliefs about the OS hazard ratio
  - Prior probability hazard ratio $= 1$ (0.90)
  - Standard deviation $\tau$ of distribution of log hazard ratio (1)
Iniparib Phase II; deaths=73, p2t=.01, prior for null = 0.9

Area = 0.71
Olaratumub Phase II; deaths=90, MLE HR=.46, prior for null = 0.9

Area = .95
# Relationship of Nominal p value for OS to Posterior Probability of Clinical Significance

<table>
<thead>
<tr>
<th># events</th>
<th>$p_{2t}$</th>
<th>$H_r ≤ .70$</th>
<th>$H_r ≤ .75$</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>.05</td>
<td>.75</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>.01</td>
<td>.88</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td>.96</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>.0001</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>50</td>
<td>.05</td>
<td>.66</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>.01</td>
<td>.83</td>
<td>.89</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td>.95</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>.0001</td>
<td>.98</td>
<td>.99</td>
</tr>
<tr>
<td>100</td>
<td>.01</td>
<td>.71</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td>.89</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>.0001</td>
<td>.96</td>
<td>.98</td>
</tr>
</tbody>
</table>
FRIENDS ANNUAL MEETING

YANPING WANG, PHD
ELI LILLY & CO.
Making provision for unexpected results in Phase 2 trials

• Suppose a phase 2 oncology trials shows unexpectedly strong results in OS

• Depending on the strength of evidence and quality of the data it may be possible to:
  • File a regulatory submission (exceptional case)
  • Expand the trial with additional patients and follow-up (very promising case)

• What minimum design-time pre-specification will permit trial expansion without undermining statistical integrity?
Basic Phase 2 Design with Provision for Expansion if Results are Promising

- Amazing
- Promising
- Poor

Sample Size

Z-statistic

Critical Value = 1.96x if trial is expanded

100 OS Events
End of Phase 2

150 OS Events
planned expansion
Case 1: Amazing result obtained. Stop trial and submit.

Critical Value = 1.96x if trial is expanded.

- Amazing
- Promising
- Poor

Sample Size:
- 100 OS Events (End of Phase 2)
- 150 OS Events (planned expansion)
Case 2: Very promising result obtained. Continue to planned expansion

Z-statistic

Sample Size

Amazing

Promising

Poor

Critical Value = 1.96x if trial is expanded

100 OS Events
End of Phase 2

150 OS Events
planned expansion
Case 3: Promising result obtained. Adapt sample size beyond planned expansion

Critical Value = 1.96x if trial is expanded

Critical Value depends on magnitude of adaptive expansion
Case 4: Poor result obtained. Stop trial without going to planned expansion