Clinical Trial Monitoring: Finding and Assessing Safety Signals

Robert Bigelow
Associate Director, Clinical Trials Statistics
Topics

- Challenges of interim safety assessment
- Progress in meeting these challenges
- Use of external data in interim safety analyses
Key Points

- Clinical trials are designed to produce data supporting a decision and optimizing ability to estimate treatment effect on primary endpoint(s)

- Usually not optimal for analysis of interim safety data

- Relatively simple statistical techniques may improve quantification of interim safety results

- Potential application of EHR data
IDMC Challenges

- Trial data may be insufficient
  - Lack of statistical power
  - Important rare events unobserved
  - Significant information external to trial

- Unreliable safety surrogates
  - Serum transaminases (alone)
  - Mild AEs not necessarily predictive of more serious ones

- Emergence of unexpected safety signal
  - New non-clinical toxicity finding
  - Safety signal from other clinical study
  - Signal from unexpected source in the trial, e.g. local IRB

- Benefit – risk considerations

- Statistical inference complicated in sequential testing
## Control of Factors Leading to Valid Inference

<table>
<thead>
<tr>
<th>Factor</th>
<th>CT primary endpoint</th>
<th>IDMC safety review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endpoint definition</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>$N \mid \Delta, \sigma, \alpha, \beta$</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Pre-specified stat analyses</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Multiplicity control</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Endpoint adjudication</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Stratification for prognostic factors</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Data ascertainment</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Data quality monitoring</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Freedom from observer bias / COI</td>
<td>•</td>
<td>• - •</td>
</tr>
<tr>
<td>Validity of surrogate endpoints</td>
<td>•</td>
<td>• - •</td>
</tr>
</tbody>
</table>

- **Good**
- **Fair**
- **Still a challenge**
Progress

- AESI - Adverse events of special interest are [ref 1]
  - “… of scientific and medical concern specific to the sponsor’s product or programme…”

- Drug induced liver injury [ref 2]
  - ‘single most specific predictor of a drug’s potential for liver toxicity is
    ✓ occurrence of cases of aminotransferase (AT) elevation + increase in serum total bilirubin, not explained by other causes … and
    ✓ increased incidence of AT in trial population, compared to controls’

- FDA DMC Guidance [ref 3]
  - Shows relationship between IDMC practices and credibility of CT results
  - Especially for safety, describes balance between judgment and quantitative analysis
# Use of External Data

Example: Comparison with published results from completed CTs

<table>
<thead>
<tr>
<th></th>
<th>Charisma</th>
<th>Plato</th>
<th>Triton</th>
<th>Ongoing Study # 1</th>
<th>Ongoing Study # 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment Control</strong></td>
<td>Clopidogrel + ASA Placebo + ASA</td>
<td>Tricagrelor Clopidogrel</td>
<td>Prasugrel Clopidogrel</td>
<td>Test Placebo</td>
<td>Test Placebo</td>
</tr>
<tr>
<td><strong>Patient Population</strong></td>
<td>Admitted to the hospital with an ACS</td>
<td>Admitted to the hospital with an ACS</td>
<td>Moderate-high risk ACS with scheduled PCI</td>
<td>Subjects with history of atherosclerotic disease</td>
<td>Admitted to hospital with moderate-high risk ACS</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>7802 + 7801 = 15,603</td>
<td>9333 + 9291 = 18,624</td>
<td>6813 + 6795 = 15,608</td>
<td>13,225+ 13,223 = 26,448</td>
<td>6423 + 6428 = 12,851</td>
</tr>
<tr>
<td><strong>Follow up</strong></td>
<td>Med. 28 months</td>
<td>12 months</td>
<td>6-15 months</td>
<td>So far: med. 20 months</td>
<td>Med. 11 months</td>
</tr>
<tr>
<td><strong>Composite of CV death, MI, stroke</strong></td>
<td>During f/u period: 6.8% vs 7.3% RR:0.93 (0.88, 1.05) P = 0.22</td>
<td>12-month 9.8% vs. 11.7% HR: 0.84 (0.77, 0.92) P &lt; 0.0001</td>
<td>15-month 9.9% vs. 12.1% HR: 0.81 (0.73,0.90) P &lt; 0.0001</td>
<td>12-month (CEC) 3.1% vs. 3.6% HR: 0.86 (0.77, 0.97) P = 0.015</td>
<td>12-month (CEC) 7.3% vs. 7.9% HR: 0.93 (0.81,1.06) P = 0.266</td>
</tr>
<tr>
<td><strong>Mortality</strong></td>
<td>4.8% vs. 4.8% RR: 0.99 (0.86, 1.14) P = 0.90</td>
<td>4.5% vs. 5.5% HR: 0.78(0.69,0.89) P&lt;0.001</td>
<td>3.0% vs. 3.2% HR: 0.95 (0.78,1.16) P = 0.64</td>
<td>1.4% vs. 1.5% HR: 0.95 (0.81, 1.11) P = 0.537</td>
<td>4.0% vs. 3.8% HR: 1.05 (0.88, 1.26) P = 0.587</td>
</tr>
<tr>
<td><strong>Bleeding</strong></td>
<td>Gusto severe 1.7% vs. 1.3% RR: 1.25 (0.97,1.61)</td>
<td>TIMI major 7.9% vs. 7.7% HR: 1.03 (0.93,1.13)</td>
<td>TIMI major or minor 5.0% vs. 3.8% HR: 1.31 (1.11, 1.56)</td>
<td>TIMI major 1.0% vs. 0.8% HR: 1.35 (1.07,1.70)</td>
<td>TIMI major 1.7% vs. 1.1% HR: 1.73 (1.26, 2.38)</td>
</tr>
<tr>
<td><strong>Fatal bleeding</strong></td>
<td>0.3% vs. 0.2% RR: 1.53 (0.83, 2.82) P = 0.17</td>
<td>0.3% vs. 0.3% HR: 0.87 (0.48, 1.59) P = 0.66</td>
<td>0.4% vs. 0.1% HR: 4.19 (1.58,11.11) P = 0.002</td>
<td>0.09% vs. 0.07% HR: 1.90 (0.92, 3.95) P = 0.084</td>
<td>0.3% vs. 0.2% HR: 1.85 (0.88, 4.06) P = 0.104</td>
</tr>
<tr>
<td><strong>ICH</strong></td>
<td>0.3% vs. 0.3% RR: 0.96 (0.55,1.65)</td>
<td>0.3% vs. 0.2% HR: 1.87 (0.98, 3.58)</td>
<td>0.3% vs. 0.3% HR: 1.12 (0.58, 2.15)</td>
<td>0.3% vs. 0.2% HR: 1.39 (0.85, 2.18)</td>
<td>0.3% vs. 0.2% HR: 1.88 (0.84, 4.23)</td>
</tr>
<tr>
<td><strong>Endpoint events prevented per 1,000 subjects treated</strong></td>
<td>4/1,000 (5 Stroke + 1MI - 5 CV death)</td>
<td>20/1,000 (9 CV death + 2 stroke + 9 MI)</td>
<td>23/1,000 (21 MI + 2 CV death)</td>
<td>7/1,000 (1 CV death + 4 MI + 2 stroke) Among fu &gt;= 1yr</td>
<td>6/1,000 (5 MI + 1 CV death) Among subjects fu &gt;= 1yr</td>
</tr>
<tr>
<td><strong>Excessive fatal bleeding per 1,000 subjects treated</strong></td>
<td>1/1,000</td>
<td>0/1,000</td>
<td>2/1,000</td>
<td>0.8/1,000</td>
<td>2/1,000</td>
</tr>
<tr>
<td><strong>Excessive bleeding per 1,000 subjects treated</strong></td>
<td>Gusto severe 4/1,000</td>
<td>TIMI major 2/1,000</td>
<td>TIMI major/minor 10/1,000</td>
<td>TIMI major/minor 7/1,000</td>
<td>TIMI major/minor 11/1,000</td>
</tr>
</tbody>
</table>

 Courtesy of Kerry Lee and Zhen Huang
Use of External Data: Bayesian Approaches

- **Use of pre-specified boundaries for AE rates** (Yao et al.)
  - For on-study AE rate ($\pi$) assume prior $\sim \beta(a,b)$
  - Update posterior as trial accrues
  - Stop if $\Pr(\pi - \pi_c | \text{prior for } \pi, \pi_c, \text{data}) \geq p$ [pre-specified threshold]
  - Applicable to studies with or without concurrent controls
  - Prior needs to allow for sensitivity to accruing data!!!

- **Meta-analysis could be effective in defining a prior accounting for study-to-study variation**
  - Still remains to show that prior is relevant for current study
  - May involve hypothesis test or measure of fit + expert clinical judgment

---

EHR: Important Adverse Events

- From EU-ADR project
- 23 Important events based on scoring for
  - trigger for drug withdrawal
  - trigger for black box warning
  - leading to emergency department visit or hospital admission
  - probability of event to be drug-related
  - likelihood of death

Table 2. List of important events in pharmacovigilance, grouped according to system/organ involved

<table>
<thead>
<tr>
<th>System/organ</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hematologic</td>
<td>Hemolytic anemia, Aplastic anemia/pancytopenia, Neutropenia, Thrombocytopenia</td>
</tr>
<tr>
<td>Cutaneous</td>
<td>Maculo-papular erythematous eruptions, Bullous eruptions (Stevens Johnson Syndrome, Lyell’s Syndrome)</td>
</tr>
<tr>
<td>Liver and gastrointestinal</td>
<td>Acute liver injury, Acute pancreatitis, Upper gastrointestinal bleeding</td>
</tr>
<tr>
<td>Cardiac and vascular</td>
<td>Acute myocardial infarction, QT prolongation, Cardiac valve fibrosis, Venous thrombosis</td>
</tr>
<tr>
<td>Neurologic</td>
<td>Convulsions, Peripheral neuropathy, Extrapyramidal disorders, Rhabdomyolysis</td>
</tr>
<tr>
<td>Psychiatric</td>
<td>Confusional state, Mood changes: depression and mania, Amnesias, Suicidal behavior/attempt</td>
</tr>
<tr>
<td>Renal</td>
<td>Acute renal failure, Anaphylactic shock</td>
</tr>
<tr>
<td>Multi-systemic</td>
<td></td>
</tr>
</tbody>
</table>

Summary

- Development of reliable toxicity predictors is valuable

- Quantification of external information
  - Frequentist thinking useful, but may fail to use relevant information
  - Bayesian approach depends on relevancy of prior data to current study

- EHR contains information that may help characterize important AEs

- IDMC members should be aware of challenges in the analysis of interim data
References


3. Guidance for Clinical Trial Sponsors, Establishment and Operation of Clinical Trial Data Monitoring Committees – March 2006


Thank You
AE Incidence Estimation in Two-Stage Design

Distribution of Estimated AE Incidence from Study of 17 Patients
Comparison of Unstaged and Two-stage Design (Stop if > 28 Patients have Events in Stage 1)
Assuming True Incidence of 30%

Mean 30.0
Std Deviation 11.1

Distribution of Estimated AE Incidence from Study of 203 Patients
Comparison of Unstaged and Two-stage Design (Stop if > 25 / 78 Patients have Events in Stage 1)
Assuming True Incidence of 30%

Mean 30.0
Std Deviation 3.2

Mean 33.6
Std Deviation 13.8

Mean 31.1
Std Deviation 4.3