Making better use of early phase safety data
Laurence Colin / Yue Li
May 10, 2019
Safety is a major cause of failed drug approvals

Table 5. Frequency of Safety, Efficacy, CMC, and Labeling Deficiencies for Drugs Failing First-Cycle Review

<table>
<thead>
<tr>
<th>Type of Deficiency</th>
<th>First-Cycle Review Failures (n = 151)</th>
<th>Delayed Approvals Following Resubmission (n = 71)</th>
<th>Drugs Never Approved During Study (n = 80)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy deficiencies only</td>
<td>48 (31.8)</td>
<td>15 (21.1)</td>
<td>33 (41.3)</td>
<td>.01</td>
</tr>
<tr>
<td>Safety and efficacy deficiencies</td>
<td>41 (27.2)</td>
<td>13 (18.3)</td>
<td>28 (35.0)</td>
<td>.03</td>
</tr>
<tr>
<td>Safety deficiencies only</td>
<td>39 (25.8)</td>
<td>24 (33.8)</td>
<td>15 (18.8)</td>
<td>.04</td>
</tr>
<tr>
<td>CMC alone</td>
<td>17 (11.3)</td>
<td>13 (18.3)</td>
<td>4 (5.0)</td>
<td>.02</td>
</tr>
<tr>
<td>Labeling alone</td>
<td>4 (2.6)</td>
<td>4 (5.6)</td>
<td>0</td>
<td>.05</td>
</tr>
<tr>
<td>CMC and labeling</td>
<td>2 (1.3)</td>
<td>2 (2.8)</td>
<td>0</td>
<td>.22</td>
</tr>
</tbody>
</table>

Three angles for safety in early drug development

• Toxicology studies in animals identify target organs

• Phase I dose-escalation studies in humans with primary goal to study safety

• Once safety risk is identified, Phase IIb studies determine a dose with optimal benefit-risk profile
Dose-escalation studies

Focus of safety analyses:
- Immediate safety of trial subjects
- Determination of safety profile of investigational compound

- Dose 1: n = 6 (+2)
- Dose 2: n = 6 (+2)
- Dose 3: n = 6 (+2)
- Dose 4: n = 6 (+2)
The approach we suggest for internal decision making

• Analyze continuous changes from baseline in laboratory parameters and relate them to drug exposure

• When a signal arises, put it in context with expected incidence without drug (‘virtual safety controls’)

Exposure-response analyses for safety

• Very successful in cardiac safety (QT prolongations)

• Controversial for liver safety (first-pass effect), however exposure-ALT relationships exist with many hepatotoxic compounds
Exposure-response analyses for liver safety

• Lumiracoxib: withdrawn from global markets in 2007 due to hepatotoxicity

• Dose-escalation study: 5 doses, 30 volunteers

\[ \Delta ALT_i = \alpha + \beta \cdot \log(AUC_i), \quad i = 1, \ldots, 30 \]

• \( \beta > 0 \) (p=0.09)
Putting signals in context with expected incidence without drug

• Examples of signals from first-in-man studies:
  • Moderate ALT elevation (>ULN) in 1/6 healthy subjects receiving active drug
  • Moderate heart rate elevations (by >20 bpm) in 2/6 healthy subjects receiving active drug
  • Marked amylase elevation (>2 ULN) in 1/6 healthy subjects receiving active drug

Question: how likely is it that these are related to drug?
Data we can use to assess relationship to drug

• 3 desirable characteristics:
  • As close as possible to a clean control group
  • With the most amount of raw data (subject characteristics, longitudinal measurements if possible)
  • Large sample

• Internal historical studies often win over real world evidence for first 2 characteristics
Example: a subject in a first-in-man study with abnormal amylase

• 1/6 healthy volunteers treated with investigational drug presents with amylase > 2 ULN

• We have raw data from 99 historical Novartis studies in healthy volunteers where placebo has been given

• From these historical data, what is our best estimate of the probability that 1/6 subject would have amylase > 2 ULN under placebo?
Novartis healthy volunteer studies with placebo

<table>
<thead>
<tr>
<th>N = 1775 subjects (99 studies)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender: Male</td>
<td>1471</td>
<td>82.9</td>
</tr>
<tr>
<td>Ethnicity: White</td>
<td>1203</td>
<td>67.8</td>
</tr>
<tr>
<td>Asian</td>
<td>287</td>
<td>16.2</td>
</tr>
<tr>
<td>Black</td>
<td>215</td>
<td>12.1</td>
</tr>
<tr>
<td>Native American</td>
<td>12</td>
<td>0.68</td>
</tr>
<tr>
<td>Other</td>
<td>58</td>
<td>3.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>median</th>
<th>IQR</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>34</td>
<td>26-44</td>
<td>18-78</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>175</td>
<td>168.5-181</td>
<td>143.8-199</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>77.9</td>
<td>69-86.4</td>
<td>47.7-116.1</td>
</tr>
</tbody>
</table>
Is 1/6 subjects with amylase>2ULN likely under placebo?
Model for probability of ALT>ULN (healthy volunteer placebo database)

• When subject characteristics, baseline ALT, number of measurements are available, more precise answers can be given

• Probability of ALT > ULN modeled on the logistic scale

• Random effect for study

• Fixed effects: baseline ALT, age, weight, number of post-baseline samples
The risk distribution is skewed and baseline is a strong predictor.

Distribution of individual predicted probabilities for ALT > ULN

Predicted probability of ALT > ULN vs. baseline ALT (U/L)
## Risk prediction for each subject (had they been taking placebo)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Baseline ALT (U/L)</th>
<th>ULN (U/L)</th>
<th>Number of post-baseline samples taken</th>
<th>Age (years)</th>
<th>Weight (kg)</th>
<th>Predicted probability ALT&gt;ULN under placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>55</td>
<td>5</td>
<td>22</td>
<td>75</td>
<td>1.4%</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>55</td>
<td>5</td>
<td>32</td>
<td>78</td>
<td>2.9%</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>55</td>
<td>5</td>
<td>47</td>
<td>70</td>
<td>6.2%</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>55</td>
<td>5</td>
<td>25</td>
<td>80</td>
<td>15.7%</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>55</td>
<td>5</td>
<td>52</td>
<td>76</td>
<td>14.4%</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>55</td>
<td>5</td>
<td>35</td>
<td>77</td>
<td>33.9%</td>
</tr>
</tbody>
</table>
Nothing unusual for liver safety in this cohort

• For this cohort, the predicted probability to observe at least one ALT > ULN is 57%
• Therefore, there is nothing unusual about observing 1/6 ALT > ULN under drug
Other applications: in-licensing evaluations

• A first-in-man study with 4/44 subjects with heart rate elevations above 100 bpm (would be expected 9.8% of the time)

• A first-in-human study with 3/90 lipase elevations > 3x ULN (would be expected 1.6% of the time)
Other applications: pediatric patients in studies

• Following the PREA legislation, we are including more pediatric patients in clinical studies

• Existing reference ranges for pediatric liver parameters are not very reliable (and known to differ by age group/sex)

• Using real world evidence (claims database), we can calculate probability that an 8-year-old would present with a given lab value in the real world
Conclusions

• Using large databases, we can help quantifying the probability that a small safety signal would have happened in the absence of drug

• Statisticians have a role to play in helping their organizations quantify the level of risk in every decision:

• ‘We estimate that the events observed in the current study would have been very unlikely to happen under placebo’

• Statisticians also have a role to play in communicating the limitations of the risk assessment
References

Thank you