Evaluating Adaptive Dose Ranging Studies:
A Report from the PhRMA Working Group

José Pinheiro, Novartis Pharmaceuticals
on behalf of the ADRS WG

Rutgers Biostatistics Day – 02/16/07
Outline

• Background, goals and scope
• Simulation study and sample results
• Conclusions
• Recommendations
Adaptive Dose Ranging Studies core WG members

- Alex Dmitrienko, Eli Lilly
- Amit Roy, BMS
- Brenda Gaydos, Eli Lilly
- Frank Bretz, Novartis
- Frank Shen, BMS
- Greg Enas, Eli Lilly
- José Pinheiro, Novartis
- Michael Krams, Pfizer

- Qing Liu, J & J
- Rick Sax, AstraZeneca
- Tom Parke, Tessella
ADRS additional WG members

- Björn Bornkamp, University of Dortmund
- Beat Neuenschwander, Novartis
- Chyi-Hung Hsu, Pfizer
- Franz König, Med. Univ. Vienna
Background

- Pharma industry pipeline problem: fewer approvals and increasing costs
- FDA Critical Path Initiative – “Innovation vs. Stagnation” white paper
- PhRMA’s response: BCG survey and report identifying key drivers of poor performance and proposing solutions
- Pharmaceutical Innovation Steering Committee (PISC) formed 10 working groups to implement BCG proposals: Rolling Dose Studies (later Adaptive Dose Ranging Studies) and Novel Adaptive Designs among them
ADRS initiative – Goals

• Investigate and develop designs and methods for efficiently learning about safety and efficacy DR profile $\Rightarrow$ benefit/risk profile

• More accurate and faster decision making on dose selection and improved labeling

• Evaluate statistical operational characteristics of alternative designs and methods to make recommendations on their use in practice

• Increase awareness about this class of designs, promoting their use, when advantageous
ADRS – Definition and Scope

- Adaptive dose-ranging designs allowing dynamic allocation of patients and possibly variable number of dose levels based on accumulating information.

- Intended to strike balance between need for additional DR information and increased costs and time-lines.

- Emphasis on modeling/estimation (learning) as opposed to hypothesis testing (confirming).

- Investigate existing and new ADRS methods via simulation.

- Evaluate potential benefits over traditional dose-ranging designs over variety of scenarios to make recommendations on practical usefulness of ADRS methods.
Simulation study: design and assumptions

• Proof-of-concept + dose finding trial, motivated by neuropathic pain indication (conclusions and recommendations can be generalized)

• Key questions: whether there is evidence of dose response and, if so, which dose level to bring to confirmatory phase and how well dose response (DR) curve is estimated

• Primary endpoint: change from baseline in VAS at Week 6 (continuous, normally distributed)

• Dose design scenarios (parallel arms):
  – 5 equally spaced doses levels 0, 2, 4, 6, 8
  – 7 unequally spaced dose levels: 0, 2, 3, 4, 5, 6, 8
  – 9 equally spaced dose levels: 0, 1, . . . , 8

• Significance level: one-sided FWER $\alpha = 0.05$

• Sample sizes: 150 and 250 patients (total)
Dose response profiles

Expected change from baseline in VAS at Week 6

-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5

Dose
Dose finding methods in simulation

- Traditional **ANOVA** based on pairwise comparisons and multiplicity adjustment (Dunnett)
- **MCP-Mod** combination of multiple comparison procedure (MCP) and modeling (Bretz, Pinheiro and Branson, 2005)
- **MTT**: novel method based on Multiple Trend Tests
- Bayesian Model Averaging: **BMA**
- Nonparametric local regression fitting: **LOCFIT**
- **GADA**: Dynamic dose allocation based on Bayesian normal dynamic linear model (Krams, Lees and Berry, 2005)
- **D-opt**: adaptive dose allocation based on D-optimality criterion
Measuring performance

- Probability of identifying dose response: $Pr(DR)$

- Probability of identifying clinical relevance and selecting a dose for confirmatory phase: $Pr(dose)$

- Dose selection
  - Distribution of selected doses (rounded to nearest integer, if continuous estimate possible)
**Dose selection performance (cont.)**

- Target dose interval – doses that produce effect within $\pm 10\%$ of target effect $\Delta$

<table>
<thead>
<tr>
<th>Model</th>
<th>Target dose</th>
<th>Target interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>actual</td>
<td>rounded</td>
</tr>
<tr>
<td>Linear</td>
<td>6.30</td>
<td>6</td>
</tr>
<tr>
<td>Logistic</td>
<td>4.96</td>
<td>5</td>
</tr>
<tr>
<td>Umbrella</td>
<td>3.24</td>
<td>3</td>
</tr>
<tr>
<td>Emax</td>
<td>2.00</td>
<td>2</td>
</tr>
<tr>
<td>Sig-Emax</td>
<td>5.06</td>
<td>5</td>
</tr>
</tbody>
</table>

- Probabilities of under-, over-, and correct interval estimation:

$$P^- = P(\hat{d}_{\text{targ}} < d_{\text{min}}), \quad P^+ = P(\hat{d}_{\text{targ}} > d_{\text{min}}),$$

$$P^\circ = 1 - (P^- + P^+)$$
Sample of Simulation Results
Probability of identifying DR

Pr(DR)

ANOVA
BMA
Dopt
GADA
MCPMod
MTT
LOCFIT

60 70 80 90 100
N = 150
linear
umbrella
logistic
Emax

60 70 80 90 100
N = 250
linear
umbrella
logistic
Emax

5 doses 7 doses 9 doses

N = 150 N = 250 N = 150 N = 250

Pr(DR)
Probability dose selection under flat DR
### Probability dose selection under active DR

<table>
<thead>
<tr>
<th>N = 250</th>
<th>N = 250</th>
<th>N = 250</th>
<th>N = 250</th>
</tr>
</thead>
<tbody>
<tr>
<td>logistic</td>
<td>umbrella</td>
<td>linear</td>
<td>Emax</td>
</tr>
<tr>
<td>LOCFIT</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>BMA</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>MTT</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>MCPMod</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>GADA</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Dopi</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>ANOVA</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N = 150</th>
<th>N = 150</th>
<th>N = 150</th>
<th>N = 150</th>
</tr>
</thead>
<tbody>
<tr>
<td>logistic</td>
<td>umbrella</td>
<td>linear</td>
<td>Emax</td>
</tr>
<tr>
<td>LOCFIT</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>BMA</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>MTT</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>MCPMod</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>GADA</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Dopi</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>ANOVA</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Pr(dose)
Probability of correct interval dose selection

Correct target interval probability (%)
Estimated dose distrib., Logistic model and N = 150

Dose selected
Estimated dose distrib., Umbrella model and N = 150
Average prediction error per dose, N = 150

Average prediction error relative to target effect (%)
Sample predicted curves: Logistic, 9 doses and N = 150
Sample predicted curves: Umbrella, 5 doses and \( N = 250 \)
Conclusions

- Detecting DR is considerably easier than estimating it.
- Current sample sizes for DF studies, based on power to detect DR, are inappropriate for dose selection and DR estimation.
- None of the methods had good performance in estimating dose in the correct target interval: maximum observed percentage of correct interval selection – 60% \(\Rightarrow\) larger \(N\) needed.
- Adaptive dose-ranging methods (i.e., ADRS) lead to gains in power to detect DR, precision to select target dose, and to estimate DR – greatest potential in the latter two.
• Model-based methods have superior performance compared to methods based on hypothesis testing

• Number of doses larger than 5 does not seem to produce significant gains (provided overall \( N \) is fixed) \( \implies \) trade-off between more detail about DR and less precision at each dose

• In practice, need to balance gains associated with adaptive dose ranging designs approach against greater methodological and operational complexity
Recommendations

- Adaptive, model-based dose-ranging designs should be used routinely in drug development, as they can lead to substantial gains in performance over traditional DF methods.

- Sample size calculations for Phase II studies should take into account desired precision of estimated target dose and possibly also estimated DR (current methods are not appropriate).

- When resulting sample size is not feasible, should consider selecting two or three doses for confirmatory phase to increase likelihood of including “correct” dose – adaptive designs could be used in confirmatory phase for greater efficiency (e.g., dropping less efficient doses earlier).
Recommendations (cont.)

- Proof-of-concept (PoC) and dose selection should be combined, when feasible, into one seamless trial.
- Early stopping rules, for both efficacy and futility, should be used when feasible to allow greater efficiency in adaptive designs – Bayesian methods are particularly well-suited for this purpose.
- Trial simulations should be used to determine appropriate sample sizes, as well as for estimating operational characteristics of designs/methods under consideration.
- Explore pre-competitive consortium for Adaptive Designs software, including ADRS.
References
