Adaptive Designs in Clinical Development:
An introduction to methods and utility

Wednesday November 9, 2016
Adaptive Designs in Clinical Development: An introduction to methods and utility

Speakers:
- **Stella Sergiopolous**, MS, MPH, Research Fellow, Tufts Center for the Study of Drug Development
- **Mark Chang**, PhD, Senior Vice President, Veristat LLC
- **Cyrus Mehta**, PhD, President, Cytel
- **Deepak L. Bhatt**, MD, MPH, Executive Director of Interventional Cardiovascular Programs, Brigham & Women’s Hospital; Professor, Harvard Medical School
- **Yi Liu**, PhD, Senior Manager, Takeda Pharmaceuticals International Co.

Moderator:
- **John Balser**, PhD, Co-founder and President, Veristat LLC
Some issues we will try to address:

- What are the main types of adaptive design?
- How often are pharmaceutical and biotech companies using adaptive approaches?
- What is the impact of adaptive designs on the cost and efficiency of clinical development programs?
- How do biotech and pharma companies make informed decisions as to whether an adaptive design would be useful in their clinical development programs?
- Should adaptive methods be used early in the clinical investigation of new therapies?
- Are adaptive designs in later phase studies different than those in early phase?
- Are there major differences in the methods used in early versus later phase studies?
- What are some examples of the use of adaptive designs?
A Landscape Analysis of Adaptive Trial Design

Stella Stergiopoulos
Research Fellow
Tufts Center for the Study of Drug Development (Tufts CSDD)

Mass Bio Talk; November 9, 2016
Agenda

• About Tufts CSDD
• Defining Adaptive Trial Design (ATD)
• Regulatory Environment
• Estimated Prevalence
• Value Proposition
About Tufts CSDD

- Center at Tufts University School of Medicine (http://csdd.tufts.edu)

- Independent, academic group focusing on drug development scientific, regulatory, economic and management policy

- Grant funded and (cross-industry) sponsored studies

- Results have informed Congress, the National Academies of Science, Foundations, Industry, Capital Market analysts and investors, Regulatory Agencies, the National Institutes of Health
TCSDD Select Publications on Protocol Design

- Stergiopoulos S, Eustace C, Stem K, Getz K. Mobile nurse services in clinical trials: usage and industry perceptions Therapeutic Innovation & Regulatory Science; Published online before print January 27, 2016. 2016.
The Decade of Optimized Protocol Design: 2011 - 2020

• Current R&D paradigm and global market conditions demand transformative solutions

• Deep pocket, solely science-driven R&D no longer an option

• Must overcome fear of changing legacy processes

• Alignment with regulatory agency interest in supporting achievement of better quality and efficiency

• Primary optimization opportunities include:
  – Traditional protocol design optimization (out of scope today)
  – Adaptive trial design
Adaptive Trial Designs Defined

- **Preplanned adaptations** – generated through the use of trial simulations and scenario planning – of one or more specified clinical trial design elements that are modified and adjusted **while the trial is underway** based on analysis of interim data.

- **Timeline Notation:**

  [Diagram with timelines and phases]

  - **Target**
    - Target discovery and validation
  - **PoC**
    - PoC clinical trials
  - **Clinical development**
    - Confirmatory phase
  - **Approval**

  - **Exploratory phase**
    - Apply biomarkers, modelling and simulation, and advanced statistical methodology
    - Demonstrate PoC and establish dose selection

  - **Confirmatory phase**
    - Apply innovative tools and clinical trial designs such as adaptive or seamless studies
    - Identify target patient population, confirm optimal dose and dosing regimen and establish the benefit/risk ratio

Main Types of Adaptive Trial Designs

<table>
<thead>
<tr>
<th>Design</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stopping Trial Early</td>
<td>Stopping clinical trial early on boundary crossing: Superiority, Futility, and/or Harm</td>
</tr>
<tr>
<td>Sample Size Re-Estimation</td>
<td>Increasing or decreasing the sample size at an interim point in the trial. Used when unclear on between-subject variance (blinded data or unblinded) a clinically meaningful effect size at which to power the trial (unblinded data).</td>
</tr>
<tr>
<td>Seamless Trials</td>
<td>Combining in one trial objectives that are typically in separate trials; i.e., learn (dose selection) and confirm (efficacy). Can either use data from final stage (operationally seamless) or combine from both phases for final analysis (inferentially seamless)</td>
</tr>
<tr>
<td>Dose-Ranging</td>
<td>Allowing for dynamic allocation of patients and possibly variable number of dose levels based on accumulating information</td>
</tr>
</tbody>
</table>

Sources:
Current Regulatory Guidance in Drug Development

2010 Guidance (CDER + CBER)

Guidance for Industry

Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the Federal Register of the notice announcing the availability of the draft guidance. Submit comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the Federal Register.

For questions regarding this draft document contact Robert O’Neill or Sue-Jan Wang at 301-796-1700, Marc Walfon at 301-796-2600 (CDER), or the Office of Communication, Outreach and Development (CBER) at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

February 2010
Clinical/Medical

2016 Reflections (CBER)
Summary of 2010 Guidance

• All planned adaptations should be specified at the design state
• Recommend simulation studies
• Two categories based on regulatory experience:
  – Well-understood (e.g. traditional group sequential design)
  – Less well-understood (but still allowed with proper rationale)
• Study design must:
  – Control for Type I error
  – Minimize:
    • Impact of statistical or operational bias on treatment effect estimates
    • Interpretability of results
• Things to look out for:
  – Timing of decision
  – Type I error control
  – Operational bias

Rationale for the TCSDD 2013 Adaptive Trial Design Study

• Growing body of knowledge in the literature

• Strong number of industry conferences and excellent response to initial TCSDD RT program

• Limited upper management and cross-functional awareness and understanding

• Few metrics and insights into the impact of adoption
TCSDD 2013 Study Methodology

• In-depth profiles of 13 top 25 companies based on interviews with senior management
  – 7:13 within biostatistics functions
  – 6:13 TA heads and clinical

• Conducted late 2012 – Early 2013 with a focus on:
  – Overall adoption and specific approaches used
  – Perceptions and barriers to adoption
  – Main areas of impact
2013 Study Adoption of ATDs

• Across the industry, simple adaptive designs are used in an estimated 20% of clinical trials
  – Early terminations due to futility appear to be relatively easy to implement and are becoming more widely adopted particularly in phase III studies
  – Although sample size re-estimations are considered by some large pharmaceutical companies as simple adaptive designs, industry-wide adoption has not yet occurred.

• Several leading companies report using simple adaptive designs more widely on 30% to up to 50% of phase II and III clinical trials

• Adoption of more sophisticated adaptive designs -- e.g., adaptive dose finding and seamless phase II/III trials – appears very low
2016 FDA Study: Prevalence of ATDs in IND/IDE Applications

Source: Lin et al 2016. N = 1,225 applications; 11.4% have adaptive design features.
2013 Study Barriers to Adoption: Primarily Internal

1. Operating concerns: Delays and disruptions to trial execution, patient participation and clinical supplies

2. Data management concerns
   - Monitoring data without introducing bias;
   - Lack of internal and contract provider expertise with ATDs;
   - Gaps in infrastructure and technology to implement sophisticated designs;
   - Limited experience and capacity among independent data monitoring committees

3. Regulatory affairs concerns: Lack of regulatory agency clarity engenders feeling of risk-taking
2013 Study: Implementation Insights

• Senior management buy-in and cross-functional support essential

• Most effectively initiated with simple adaptive design approaches

• Routinely measured and publicized impact values (e.g., cost and time savings)

• Assign internal adaptive trial design experts to educate and collaborate with clinical teams
TCSDD 2013 Study: ATD Impact

• Already offering cross-functional teams **new insights** into study design **through** scenario planning and trial **simulation** – ‘stress test’ protocols during development planning

• Most important impact is on success rates

• Tufts CSDD modeling of savings from early terminations due to futility:

<table>
<thead>
<tr>
<th></th>
<th>Mid-Size Company</th>
<th>Large Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Cost</td>
<td>$5 - $10MM</td>
<td>$20 - $50 MM</td>
</tr>
<tr>
<td>Indirect Cost</td>
<td>$15 - $35 MM</td>
<td>$50-$100MM</td>
</tr>
</tbody>
</table>

• Global top 20 pharmaceutical company reports **saving $70 million annually**
  – PII: Simple adaptive dose-finding
  – PIII: sample size adjustment and futility stopping
Summary of 2013 Study Findings

Save development timeline + Saving costs by stopping early / assessing sample size

“Every year, we add up how much money we saved, and it’s always in the tens of millions across the entire portfolio” --- Large Pharmaceutical Company
Thank you!

Stella Stergiopoulos
Research Fellow
Tufts Center for the Study of Drug Development (Tufts CSDD)
Stella.Stergiopoulos@tufts.edu
Early Phase Adaptive Trial Designs: Avoid Strategic Mistakes and Increase overall Probability of Success

Mark Chang, Ph.D.
Sr. Vice-President
Strategic Statistical Consulting

VERISTAT
Coverage

- Integrated sequential decision process in drug development
- Importance of right strategies in designing early phase trials
- How to design early phase trial when facing multiple challenges:
  - Limited information, heterogeneity of patient population, high-cost and financial constraints, timing, competing landscape.
Clinical Development Success Rates

- Success rate from Phase I for approval was 9.6%, and 11.9% for all indications outside of Oncology.
- Rare disease programs and programs with biomarkers had higher success rates at each phase of development than overall.
- Hematology had the highest likelihood of approval (LOA) from Phase I (26.1%) and Oncology had the lowest (5.1%).
- Phase-transition probabilities can be modelled using Markov Chain
- Presently 10% trials featuring adaptive designs, the FDA guidance encourages more adaptive designs.

Source: Biotechnology Innovation Organization (BIO), 2016
Likelihood of Approval (LoA) from Phase I

Source: Biotechnology Innovation Organization (BIO), 2016
Markov Process for Modeling Success Rates of Various Phases of Clinical Trials

Transition Probability Matrix:

Data Source: Biotechnology Innovation Organization (BIO), 2016
Traditional Phase-isolated Decision Process

Take the action (design) that maximizes the Power at Trial level.
Clinical Development Plan with Markov Decision Process

- **Markov chain** = average rate of success in clinical trials, but fail to recognize the difference and the effect of actions
- **A Markov decision process** = Markov chain + actions + rewards.

Mission: Take the actions to maximize the total expected future reward
Success Dependent on Multifactors - Solutions

1. *Unknown treatment effect* – Bayesian prior distribution of parameter
2. **Decisions** – Adaptive trial designs
3. **Probabilistic outcomes** – Simulations
4. **Interrelationship among different phases** – Markov Chain Decision Process with adaptive design to maximize the forward rewards at anytime point.
Power with Uncertainty of Treatment Effect (Prior)

- Treatment difference is a fixed but unknown value
- Prior response rate = 10%, 20%, or 30% with 1/3 probability each.
- Power = 80% based on n = 784, average effect size =20%, or
- From Bayesian, Power = (0.29+0.80+0.99)/3 = 0.69, uncertainty of treatment effect reduce the success rate.

<table>
<thead>
<tr>
<th>Effect size</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>0.29</td>
<td>0.80</td>
<td>0.99</td>
</tr>
</tbody>
</table>
How Adaptive Design can Help

- Group sequential design – potentially cut cost and shorten the time to market
- Sample-size reestimation design – protect power when the drug effect has great uncertainty
- Difficult to choose target population? – Biomarker enrichment design.
- Uncertainty about the optimal dose/regimen? – adaptive-dose finding (drop-arm and add-arm design)
Monte Carlo or Clinical Trial Simulation

- Clinical trial simulation (CTS) is a process that mimics clinical trials using computer programs.
- CTS is particularly important in adaptive designs for several reasons:
  - Intuitive approach with minimal cost, easy to implement, can be done in a short time
  - Easy to mimic complex situations
  - Provide comprehensive outputs, operating characteristics
  - Can be used to monitor trials, predict outcomes, anticipate problems, and suggest early remedies before it is too late

Example of Utility Function

Utility depends on: Associated costs, statistical significance, clinical significance, commercial viable

Define Utility:

\[ U(\hat{\delta}, n) = H(\alpha - p) \left( \hat{\delta} - \delta_{\text{min}} \right) (c_1 \hat{\delta}) + c_0 - c_2 n \]

\[ H(x) = \begin{cases} 
1 & \text{if } x \geq 0, \\
0 & \text{if } x < 0 
\end{cases} \]

Constants: \( c_0, c_1, c_2 \). Final observed treatment effect: \( \hat{\delta} \).
Thinking Strategically

- What is the value proposition?
  - indication, target population, differential values-more efficacy, safer, administrative convenience, cheaper?, value differentiation
  - Drug label

- How is the competing landscape?
  - Marketing intelligence focus drug on market
  - Regulatory intelligence focus on on-going and future trials by our competitors

- What options are presented?
  - Target patient population size? Orphan drug designation? A companion or complementary diagnostics required?
  - US only, sequential (ROW after US) or parallel (MRCT) approach,
  - National or international exhaustion, compulsory license and parallel importation

- Net Present Value (NPV)
  - Prior info uncertainty, outcome in terms of NPV, outcome uncertainty
  - Pipeline and fallback plans
Plan Strategically

- The Science:
  - Mechanism of Action, how much evidence supported
  - Key opinion leaders (KOLs)
  - Expected efficacy and safety profile, and therapeutic widow of the NCE

- Regulatory:
  - Does the primary endpoint well established (gold standard) or require clinical outcomes to support for efficacy approval?
  - Special protocol assessment (SPA) required?
  - EU Central/non-central filing?

- Operation:
  - Patient recruitment, drug stability, storage, packaging/blinding, shipment, and randomization

- Commerce and Marketing:
  - What are the clinical and commercial values of the second endpoints?
  - Drug branding, and pricing strategy

- Reimbursement:
  - Health Plan /payer,
Strategically Mitigate Risks and Increase Success- Adaptive Design

- Successful program starts at early phases
- Be aware the gap between reality and wishful thinking
  - 40% Phase-III success rate vs. over 90% power
- Robust adaptive design against inaccurate assumptions about, target population, effect size and variability, sample size, endpoint…
- Value versus probability of success
  - Superiority versus non-inferiority
  - Biomarker-specific population versus broader target population
  - Combine Noninferiority, superiority trials, biomarker-enrichment design
Questions to Ask at the Program Level

- What are the overall efficacy and safety profile?
- What is the consequence of failure in a trial?
- Any exist or potential competitors?
- Seek Expedited Programs - accelerated approval?
- What is the financial constraints?
- Dose regiment?
- What are the requirements for PK/PD studies
- Classical or adaptive approach?
- Minimal safety requirement in sample size for the indication?
- Can the primary endpoint be measured early at the interim analysis, if not, is a biomarker available at IA?
Questions to Ask at Trial Level

- Co-primary endpoints? What is multiplicity strategy?
- Superiority/non-inferiority position? NI margin justification (any precedence?)
- Phase-combined trials?
- Adaptive dose-finding, group sequential design, sample-size re-estimation, adaptive randomization, and/or biomarker enrichment design? How many interim analyses are needed?
- Evaluation matrix and utility function?

Develop simulation program and specify simulation scenarios iteratively to determine the optimal CDP and trial designs.
Summary and Discussion

- Trial Phase-transition probabilities inform the average success rate for different indications
- The success of clinical development for an individual NME depends on the decisions made in the process
- Early phase trials using adaptive design play a critical role in the success of the entire clinical development program
- Design early phase to achieve the global development optimal.
References

Adaptive Designs for Confirmatory Trials
Two Case Studies

MassBio Adaptive Design Forum
9 November 2016

Cyrus Mehta, Ph.D., Cytel Inc and Harvard T.H.Chan
and
Deepak L. Bhatt, M.D., M.P.H., BWH and Harvard Medical School
Two case studies featuring adaptive confirmatory trials

1. The VALOR trial of acute myeloid leukemia
   • motivation
   • promising zone design
   • operating characteristics
   • strategic impact

2. Enrichment design for a targeted therapy
   • motivation
   • the MAMS approach
   • power comparison: MAMS vs conventional design
Base case study design

VALOR study – Sunesis Pharmaceuticals

Double-blind RCT of vosaroxin in relapsed / refractory AML

* After cycle 1, all subsequent cycles at 70 mg/m² vosaroxin on days 1 and 4

NCT01191801; figure taken from poster of Ravandi F. et al. (ASCO, 2012): http://meetinglibrary.asco.org/content/99304-114
Strategic considerations

- **Fixed sample size design assuming Hazard Ratio (HR) = 0.71 has 90% power** (450 patients accrued over 24 mo. and 375 events observed with 6 mo. follow-up)

- **But, if HR = 0.77, power drops to 70%**
  - 90% power at that HR would require >1.6x more patients

- **Sunesis wanted to avoid incurring high cost up-front unless assumption of HR = 0.71 turned out to be optimistic**

**Could adaptive design reduce up-front cost?**
Could it also make opportunity attractive to investors?

*Cytel analysis*
POS and efficacy at fixed sample size

Lower-than-expected efficacy yields lower POS (at same sample size)

Cytel analysis
“Buying POS” by increasing sample size

When efficacy is lower than expected, increasing sample size can boost POS

Cytel analysis
Adaptive design: Interim sample size re-assessment

Transparent, pre-specified plan to increase sample size only if interim analysis was in “promising zone”

NCT01191801; figure taken from poster of Ravandi F. et al. (ASCO, 2012): http://meetinglibrary.asco.org/content/99304-114
Operating Characteristics

### 1. Under Pessimistic Scenario, HR = 0.77 (10,000 simulations)

<table>
<thead>
<tr>
<th>Zone</th>
<th>P(Zone)</th>
<th>Power</th>
<th>Duration (months)</th>
<th>SampSize</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NonAdpt</td>
<td>Adapt</td>
<td>NonAdpt</td>
</tr>
<tr>
<td>Unf</td>
<td>25%</td>
<td>33%</td>
<td>35%</td>
<td>28</td>
</tr>
<tr>
<td>Prom</td>
<td>34%</td>
<td>71%</td>
<td>90%</td>
<td>29</td>
</tr>
<tr>
<td>Fav</td>
<td>41%</td>
<td>95%</td>
<td>95%</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>—</td>
<td>71%</td>
<td>78%</td>
<td>28</td>
</tr>
</tbody>
</table>

### 2. Under Optimistic Scenario, HR = 0.71 (10,000 simulations)

<table>
<thead>
<tr>
<th>Zone</th>
<th>P(Zone)</th>
<th>Power</th>
<th>Duration</th>
<th>SampSize</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NonAdpt</td>
<td>Adapt</td>
<td>NonAdpt</td>
</tr>
<tr>
<td>Unf</td>
<td>12%</td>
<td>57%</td>
<td>53%</td>
<td>29</td>
</tr>
<tr>
<td>Prom</td>
<td>28%</td>
<td>87%</td>
<td>99%</td>
<td>30</td>
</tr>
<tr>
<td>Fav</td>
<td>60%</td>
<td>99%</td>
<td>98%</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td>—</td>
<td>90%</td>
<td>93%</td>
<td>27</td>
</tr>
</tbody>
</table>
Comments from partners

“Sunesis’ use of an adaptive trial design offers us an opportunity to invest in this promising biopharmaceutical product candidate on terms that are a win-win for both Sunesis and Royalty Pharma:
Sunesis gains access to a flexible, novel financing structure and we are able to invest in vosaroxin at a time when we believe its likelihood of commercial success will be high.”

– Pablo Legorreta, CEO, Royalty Pharma

“The innovative yet practical design provided multiple favorable scenarios that allowed us to proceed with our pivotal Valor study … It is difficult to imagine going forward with traditional methods alone.”

– Steven Ketchum, Sr. VP R&D, Sunesis Pharmaceuticals

2 S. Ketchum, personal communication
• Create an SSR design and plot its power curve vs $\delta$
• For each $\delta$, compute its expected sample size
• Plot the power curve of corresponding GSD having the same expected sample size
• The power curves are practically identical. So what extra benefit does SSR design confer?
The Benefit is in the Conditional Power

(a) \( Z \) vs Sample Size and CP for Adaptive Design

- \( \delta = 0.3 \)
- \( \delta = 0.35 \)
- \( \delta = 0.4 \)
- Sample Size

Promising Zone

MassBio Adaptive Design Forum
GSD and AGSD have different profiles for their conditional power curves in the promising zone.

But when you take a weighted average over all values of the interim Z, the two methods have the same unconditional power curves.
Can We Do Randomized Clinical Trials Better?

Deepak L. Bhatt, MD, MPH

Executive Director of Interventional Cardiovascular Programs, BWH Heart and Vascular Center
Professor of Medicine, Harvard Medical School
Participation in a National HF Registry vs. Any Randomized Clinical Trial

Patients enrolled in RCT <0.2%

Patients not in RCT >99.8%

ADHERE, Acute Decompensated Heart Failure National Registry.
The incremental cost of the TASTE trial was $300,000, or $50 for each participant who underwent randomization.
THE CHANGING FACE OF CLINICAL TRIALS
Jeffrey M. Drazen, M.D., David P. Harrington, Ph.D., John J.V. McMurray, M.D., James H. Ware, Ph.D., and Janet Woodcock, M.D., Editors

Adaptive Designs for Clinical Trials
Deepak L. Bhatt, M.D., M.P.H., and Cyrus Mehta, Ph.D.
Adaptive Features of a Trial That Uses Sample-Size Reestimation

Adaptive Design of a Cardiovascular-Outcome Trial with Zones for Decision Making Regarding Superiority

Schematic Representation of an Adaptive Two-Stage Population-Enrichment Design

Concerns Among Physicians?

- Isn’t it cheating to look at the results as the trial is ongoing?
  - Is it ethical?
    - Shouldn’t there be a huge statistical penalty?
- Will this approach work for a rapidly enrolling, event driven trial?
  - Low event rates at time of stopping enrollment
- Sponsor may be happy if trial ends early due to cost, but not the physician
  - Secondary endpoints, subgroup analyses underpowered
Thank You!

Deepak L. Bhatt, MD, MPH
Executive Director of Interventional Cardiovascular Programs,
BWH Heart & Vascular Center
Professor of Medicine,
Harvard Medical School
1 (857) 307-1992
dlbhattmd@post.harvard.edu
• Population enrichment design in oncology
• PD-1 is a protein and cell-surface receptor that down-regulates the immune system
• A new class of drugs (pembrolizumab and nivolumab) block PD-1 and activate the immune system
• Many tumors (eg NSCLC) express PD-L1, an immunosuppressive ligand that might enhance the mechanism of action of the pembrolizumab and nivolumab
CheckMate 57: OS by Biomarker Status

Borghaei et al. [2015], supplementary materials
Stratified Randomization

MAMS Design: We will make two comparisons of Nivolumab to a common control arm

- $Z_1$ compares all Nivolumab vs all Docetaxel (test the Full Population)
- $Z_2$ only the PDL1-high subgroup vs all Docetaxel (test the Subgroup)
MAMS design with two comparisons

![Graph showing stopping boundaries vs information fraction](image)
MAMS design with two

Stopping Boundaries

Information Fraction

MassBio Adaptive Design Forum
MAMS design with two

Stopping Boundaries vs Information Fraction

-1 0 1 2 3 4 5

Information Fraction

Stopping Boundaries

a1, a2, a3, a4
b1, b2, b3, b4

full, sub
MAMS design with two comparisons

The graph shows stopping boundaries for Information Fraction. The plot includes two lines:
- Red line for full design
- Blue line for sub design

The points labeled a1, a2, a3, b1, b2, and b3 are marked on the graph, indicating critical values for stopping rules in adaptive design.
MAMS design with two
HR of PDL1(high) subgroup = 0.4

Disjunctive Power

Average Total Sample Size

Hazard Ratio of PD-L1(lo)
HR of PD-L1(high)=0.4
HR of PDL1(high) subgroup = 0.5

Disjunctive Power

Average Total Sample Size

Hazard Ratio of PD-L1(lo)
HR of PD-L1(high)=0.5

Hazard Ratio of PD-L1(lo)
HR of PD-L1(high)=0.5
HR of PDL1(high) subgroup = 0.6

Disjunctive Power

Average Total Sample Size

Hazard Ratio of PD-L1(low) = 0.6
HR of PD-L1(high) = 0.6
HR of PDL1(high) subgroup = 0.7

Disjunctive Power

Average Total Sample Size

Hazard Ratio of PD-L1(lo)
HR of PD-L1(high)=0.7

MAMS
Conventional

Hazard Ratio of PD-L1(lo)
HR of PD-L1(high)=0.7

MAMS
Conventional
Summary of Results

• MAMS has greater power and smaller expected sample size compared to conventional if PDL1 biomarker is predictive
• The greater the differential benefit from PDL1(high) in patients treated with Nivolumab, the more MAMS dominates
• But PDL1(high) and PDL1(low) patients benefit more or less equally on Nivolumab therapy, then the conventional design is better
• Choice of MAMS or conventional depends on prior belief that the novel therapy benefits on subgroup more than the other
Implementing Adaptive Designs at Takeda

Yi Liu PhD, Sr Manager
Global Statistics
Takeda Pharmaceuticals Inc Co.

Mingxiu Hu, Zhaoyang Teng, Ken Nip, Hui-min Lin

Nov 9, 2016
MassBio Adaptive Design Forum
Implementation of different Adaptive Designs

• Pivotal
  – Sample size re-estimation (SSR) designs
    • Multiple doses
    • Multiple endpoints
  – Seamless phase II/III designs
    • Adding or dropping treatment arm
    • Combining phase II and phase III
  – Adaptation on study endpoint

• Non-pivotal
  – Adaptation on patient subgroups
  – Adaptive randomization (Umbrella trial)
Case 1: Seamless phase II/III Design - dropping arms

- **Primary endpoint:** percent change from baseline at 6 months
- **Slower enrollment for 1st stage**
  - 1st stage: 6 pts/month
  - 2nd stage: 18 pts/month
- **Pause enrollment after 60th pt**
- **IA timing**
  - Enroll the 60th pt + follow up 12 wks
    - 6 mo endpt for 7/8 pts per arm
    - 12 wk endpt for 11/12 pts per arm
  - Assume 3 mos for data cleaning, data, TLF generation and interim decision making
IA Decision Rule

• One option is to calculate conditional power (CP) separately for each arm comparing to the placebo
  – If all $CP_L$, $CP_M$, $CP_H < 30\%$, stop for futility
  – If top one CP is 30\% higher than top two CP, then pick only the top one dose, and do sample size recalculation
  – Otherwise, select doses with top 2 CPs
    • sample size recalculation based on CP for the selected doses and use the smaller sample size among them.

• Alternative approaches:
  – Dose response modeling including MCP-Mod
Case 2: SSR with multiple doses and multiple endpoints

- Drug is already approved by regulatory agency with two doses L and H
- This pivotal trial is to evaluate the efficacy and safety of the new formulation (NF) comparing to the old formulation (OF) for both doses
  - Phase 3 multicenter, randomized DB, placebo controlled
  - Treatment Arms (5):
    - L NF, L OF, H NF, H OF, placebo (PBO)
  - One primary endpoint (denoted as P)
  - Two key secondary endpoints (denoted as S1 and S2)
  - All endpoints are measured at month 3
Hypothesis Testing

- 12 comparisons:
  - For P (L NF vs PBO, H NF vs PBO)
  - For S1 (L NF vs PBO, H NF vs PBO)
  - For P (L [NF vs OF], H [NF vs OF])
  - For S2 (L [NF vs OF], H [NF vs OF])
  - For S1 (L [NF vs OF], H [NF vs OF])
  - For S2 (L NF vs PBO, H NF vs PBO)

- Control type I error via Holm’s method for the two comparisons within each step.
- The tests will be performed in the sequential order and only proceed to the next level if both comparisons are significant.

Primary Interest and used for SSR
SSR rule and IA timing:

- Adaptation based on the Conditional Power (CP) for the Primary Endpoint comparing NF and OF for dose L and H

- Adaptation rule
  - Study starts with 350 per arm
  - Step-wise increase with 25 patients (per arm) each step depending on IA results
  - Maximum study size 500 patients per arm (would be the study size for a fixed design)

- Conducts IA for adaptation decision when 200 subjects per group have been enrolled and completed 3-month assessments
  - Allow sufficient time for efficacy readout before enrolling 350 pts per arm so that no enrollment hold is needed
Adaptation Rule - Example

- Conditional power
- Sample size increment

- # of patients
  - 0.00
  - 0.05
  - 0.10
  - 0.15
  - 0.20
  - 0.25
  - 0.30

- Unfavorable
- Promising
- Favorable

Graph showing conditional power and patient increments.
Study Outcome

Outcome:

- Interim analysis was conducted in 2015
- Results indicated that only 350 patients (minimum sample size) per arm are needed
- Saved 750 patients in total and $15 million in cost assuming $20,000 per patient
- Shortened development time
- FA was conducted in 2016 and met statistical significance of the primary comparison of interest
Case 3: SSR to shorten development timeline

- **Study setting**
  - Two primary endpoints
    - Response Rate/PFS
      - Accelerated approval
      - “surrogate” endpoint
    - PFS/OS
      - Full/regular approval
      - longer time to mature
  - Two analysis time points
    - ORR test first at IA
    - PFS test at FA with the possibility of early stopping at IA for overwhelming efficacy
Why adaptive

• Rare disease and slow enrollment
  – Difficult to run a randomized trial with time-to-event endpoint, unless use very aggressive assumption on effect size
• Limited knowledge on effect size
  – Need an IA to re-estimate the sample size more accurately
Study Design

Design

Type I error control:
- Two endpoints: Alpha exhaustive approach
- Two analysis time points: alpha-spending
- Adaptation at IA: weighted combination test for FA

Interim Analysis

Final Analysis

ORR

Sample Size Re-estimation

PFS

PFS
Alpha Exhaustive Design

Interim Analysis

ORR

α = 0.0125

ORR significant

Sample Size Re-estimation

PFS

α = 0.0015

ORR not significant

PFS

α = 0.0004

PFS not significant

Final Analysis

α = 0.0245

PFS

α = 0.0125

Between Two Primary Endpoints

ORR

α = 0.0125

Not significant

PFS

α = 0.0125

significant

α = 0.0125 + 0.0125 = 0.025

significant
Case 4: Combining phase 2 and 3 in oncology trials

• Traditional separate phase 2 and phase 3 designs
  – Comprehensive evaluation of efficacy and safety of Phase 2 results to make a go/no-go decision for phase 3
  – Phase 3 with primary endpoints EFS and OS
    • EFS significance is required before testing OS
    • EFS significance will trigger Accelerated approval, and trial will continue to follow OS
    • OS significance will trigger full approval
      – Can add event re-estimation for OS to increase probability of success

• Design to combine phase 2 and phase 3
  – Add an interim for phase 2 portion to further accelerate development timeline if EFS shows extremely good results
Comparison: Phase 2/3 vs. separate Phase 2 & 3

Ph 2 Enrolment

Ph 3 Enrolment

Follow Up

DCO for FA of EFS @31 m

FA of EFS; OS event re-estimation

1st GNG decision made
2nd GNG decision made if 1st is not "Go"

HR_{1EFS} <= 0.7
HR_{2EFS} <= 0.8

Ph 2 Enrolment

Ph 2 Enrl

Final No-Go

Ph 3 Enrl

Follow Up

DCO for FA of EFS @33.7 m

DCO @18.3m
GNG @21.3m

DCO for FA of EFS @44.3 m

Final Analysis (OS) @ TBD

0 6 12 18 24 30 36 42 48 54 60

15.5 m 40.3 m
Operational Challenges: Controlling Bias

• Adaptation rule
  – Set up by design statistician who has no direct involvement in any part of study conduct
  – Symmetric stepwise function to prevent potential back-calculation of IA results
  – Only disclosed in the IDMC charter appendix with limited access by design statistician, head of biostatistics, IDMC

• Interim results
  – The interim analysis will be conducted by the independent statistical center (ISC) and presented for review to the IDMC during closed session.
  – Only the ISC and the IDMC will be unblinded to the interim results while the study is ongoing.
  – Sponsor remains blinded to the interim results until completion of the study

• Communication of final sample size
  – IDMC will recommend to the sponsor executive committee the final sample size.
  – The sponsor executive committee will inform the project team of the new sample size.
Thank You!

Takeda Pharmaceuticals International Co.
Upcoming MassBio Forums

**Nov 15:** Employment Contracts; HR & L&R

**Nov 30:** What is the Value of Value Frameworks in Making Healthcare Decisions?

**Dec 7:** The Intersection of Environmental Health & Safety and Manufacturing; SEF

**Dec 13:** The Prescription Drug Abuse Epidemic: Finding Solutions with New Technologies in Formulation & Delivery Science; FDD