Early Economic Evaluation of the Effectiveness of Pharmaceutical Development Candidates in Subpopulations: A Role for Clinical Trial Modeling and Simulation

ISPOR 17th Annual International Meeting
6 June 2012
Washington, DC

Agenda

- Introduction
  - Bob Leipold, PhD (United BioSource Corporation)
    Research Scientist, Health Economics

- The Importance of Sub-Populations in Economic Evaluations
  - Ananth Kadambi, PhD (United BioSource Corporation)
    Director, Modeling and Simulation Development, Health Economics

- A Role for Clinical Trial Simulations
  - Lyn Powell, PhD (Entelos, Inc.)
    Senior Research Fellow, R&D; Vice President, Strategic Relationships

- Opportunities and Challenges: A Reality Check
  - Christy Chuang-Stein (Pfizer, Inc.)
    Vice President, Statistical Research and Consulting Center

- Questions & Answers
  - Bob Leipold
Introduction

- Consideration of sub-populations is essential in economic evaluations.
  - Disease risks and treatment benefits are not uniformly distributed through the population.
  - Costs, both monetary and health (i.e., adverse events), are often concentrated in a small portion of the population.

- Clinical trial simulations can be useful for sub-population analyses.
  - It is easier to explore treatment implications in sub-populations via simulation than via clinical trials.
    » Simulations can be used to design and support adaptive clinical trials.
  - Simulations of a wide range of virtual patients can help identify characteristics that discriminate one sub-population from another.

- Sub-populations are important to the pharmaceutical industry, but problematic.
  - Prior versus post-hoc identification of sub-groups; physiological basis for differentiation
  - Better outcomes, but at what cost?

Modeling and simulation (M&S) is used to support pharmaceutical decision making all along the pipeline.

- Molecular modeling and quantitative structure-activity relationship (QSAR) analysis for lead compound synthesis and optimization

- Pharmacokinetic (PK) and pharmacodynamic (PD) modeling
  - Analysis of preclinical data
  - Extrapolation of preclinical results to humans
  - Interpretation and extrapolation of early clinical data
  - Prediction of safety and efficacy implications of dosing decisions

- Disease modeling
  - Identification and prioritization of new targets
  - Understand drug mechanisms of action
  - Identification of candidate biomarkers
  - Exploration of effects of patient variability

- Clinical trial simulations

- Economic (cost-effectiveness) analysis
Cost effectiveness is an important consideration, but it is typically deferred until later in development.

- Gain additional confidence that a drug has a chance to make it to market before investing in economic analysis
- Gather more and better efficacy data
- Gather more and better adverse-event data
- Explore the full range of possible indications and applications

BUT

- Early economic modeling could help set priorities, define data collection requirements, and establish necessary performance criteria.
  - Add economic viability as a criterion when selecting among development candidates
  - Establish what data are required to support an economic value argument
  - Establish what efficacy, adverse event profile, and price are required for an economically viable drug

Identification and analysis of subpopulations is an important part of pharmaceutical development.

- Subpopulation analysis is particularly important for indications for which standard-of-care therapeutics already have good clinical efficacy across broad patient populations.

- Identify patient populations that are especially susceptible or resistant to the efficacious and/or adverse effects of treatment

- Subpopulation analysis is an important step on the path to approval, but these analyses should begin early to avoid regulatory issues
  - Example: FDA/NICE pull approval for Avastin for treatment of metastatic breast cancer due to lack of “totality” of evidence showing safety and efficacy for this sub-population
Consideration of sub-populations can be important in economic evaluations.

- The burdens of disease and health care costs are not uniformly distributed through the population.
- Changes in health outcomes and consequent health care costs due to treatment will also be not uniformly distributed.
- Consideration of sub-populations can provide:
  - Better estimates of population outcomes
  - Identification of sub-populations that benefit most (and least!) from treatment

A major hurdle for early economic modeling of sub-populations is the lack of relevant human data.

- Estimation of real-world patient outcomes using available human trial data is difficult using traditional methods.
  - For early economic analyses, typically only early trial data (Phase I and II) are available.
    - Small enrollment
    - Measure outcomes not relevant to real-world efficacy
  - For drugs in the peri-approval phase, Phase III data may be available but are likely not representative of post-approval target population(s).
  - Statistical analyses and PK/PD modeling can be used to make estimates, but extrapolation outside the population for which data are available is imprecise.
- Relevant economic data (e.g., HRQoL/utilities) may not be available for novel therapeutics.
  - Reasonable estimates can often be made for most economic parameters from marketed comparator therapies or data from similar indications.
Clinical trial simulation (CTS) can provide insight into both efficacy and cost effectiveness, even with limited data.

- CTS refers to computer modeling approaches that replicate actual human trials using predictive equations and/or “virtual subjects”
  - CTS is relatively fast and inexpensive compared to the cost of actual alternative trial scenarios.

- Key questions CTS can address
  - Are there specific sub-populations for which a new therapeutic is cost-effective, enabling targeted trial design, adaptive trial design, or a more focused drug development strategy?
  - How relevant are effectiveness predictions derived from clinical trial populations in the context of real-world patients?
  - What efficacy is needed in a clinical or post-market population for a new therapeutic to be useful from a cost-effectiveness perspective?
  - Where are there key data gaps that should be filled?
  - What is the quantitative impact of data gaps on predictions of effectiveness?

Example Results: CTS Applied to Novel Therapeutic for Management of Chronic Pain

- Predictive equations were derived from Phase III trials of a compound in development and used to simulate a post-market (“pragmatic”) trial.

- Different trial designs were simulated, varying the number of patients, duration, inclusion / exclusion criteria, etc.

- Each simulation produced a dataset of outcomes from which statistical measures could be derived.
  - Difference in time of successful management of pain symptoms between treatment and control
  - Difference in cost-of-care accruals
Cost-Effectiveness Predictions: Effect of Sub-population Variations

1. Dominant
2. Cost-Effective

Incremental cost ($)
Incremental days in success

Effect of Variation of Patient Inclusion Criteria on Likelihood of Trial Success

Inclusion criteria varied in CTS simulations included*:
  • Prior use of opioid for pain management
  • Presence or absence of radicular pain
  • History of depression

- Early knowledge of the likelihood of trial success could enable focused development targeted to sub-populations of interest
- Predictions of efficacy can be used to inform early economic models and fill data gaps

*Results of a subset of subpopulations evaluated are shown
Summary

- Cost effectiveness in sub-populations may be critical to address in early economic evaluations

- Modeling methodologies may provide insights into effectiveness in sub-populations even in the absence of real-world data
  - Enable better-informed decision making during development

- The following presentation provides examples of one modeling approach that could be used to successfully fill data gaps related to clinical efficacy in subpopulations that make early economic modeling challenging
  - Many companies and academic institutions offer such approaches

MECHANISTIC MATHEMATICAL MODELING OF PHYSIOLOGY AND CLINICAL TRIAL SIMULATIONS

Lyn Powell, Ph.D.

R&D Entelos
Why is drug development so hard?

- The fundamental technical challenges are:
  - Understanding how a therapeutic candidate works (mechanism)
  - Understanding how different individuals will respond to that mechanism

- Individual patient responses vary due both to genetics and environment
  - Not all drugs work in all people
  - Some drugs may be dangerous for some people

- Physiological mechanistic modeling of both therapeutic target mechanism(s) and the variability in individual response due to differences in underlying disease etiology, has the potential to make therapeutic R&D programs;
  - Faster to execute, therefore, cheaper
  - Identify responders and non-responders
“Top-down” models capture clinical outcomes by simulating key biology

Examples of the mechanisms and endpoints necessary to model CV disease physiology

Cardiovascular PhysioLab platform dynamically integrates measures of cholesterol metabolism, endothelial function and platelet activation effects on intimal tissue remodeling and CHD risk

Relevant Clinical Outputs:
- Circulating lipid levels
- Lipoprotein particle size distributions
- Plaque burden e.g. IMT, PAA
- Plaque stability e.g. vessel stress, foam cell density, lipid core thickness
- Flow-mediated dilatation (FMD)
- Coronary heart disease risk (CHD risk)

Mechanistically distinct patient types:
- Hyperlipidemic
- CAD and ACS
- T2D
- Smokers

Alternate virtual patients (VPs) represent full range of mechanistic and phenotypic variability

VPs must demonstrate appropriate responses to multiple therapies:
- HMGCoA reductase inhibitors
- CETP inhibitors
- NPC1L1 inhibitors
- Vytorin
- Nicotinic acid (niacin)
- PPARδ agonists
- Reconstituted HDL/apo AI Milano
- Anti-platelet agents (aspirin, clopidogrel)
- PCSK9 inhibitors
- PPARγ agonists

Examples of the mechanisms and endpoints necessary to model CV disease physiology
Virtual patients capture diversity in phenotype due to variability in underlying mechanisms

- Virtual patients represent biological variability inherent in the heterogeneous human population
- All virtual patients should reproduce multiple experimental data sets
  - Baseline phenotypical measures
  - Response to standard interventions

Variability in Clinical Phenotypes
- Lipoprotein profiles
- Differential statin response
- Plaque size/composition

Diversity in Underlying Mechanisms
- VLDL secretion rate
- LPL activity
- apoB clearance
- CETP activity
- ABCA1 activity
- Inflammation
- Insulin resistance

Mechanistic diversity is achieved by simultaneously exploring multiple parameters
- Each unique set of parameter values defines a virtual patient
- Parameters can be explored over a feasible range defined by the literature
Mechanistic diversity drives phenotypic diversity at baseline and on therapies within a VP cohort.

Baseline measures vs. Response to simulated statin:
- Lipid values (mg/dL):
  - TC: Total Cholesterol
  - TG: Triglycerides
  - LDL-C: Low-Density Lipoprotein Cholesterol
  - HDL-C: High-Density Lipoprotein Cholesterol
  - Percent change

**Individual VPs, VPs combined into cohorts and virtual populations (Vpops) have different applications**

- **Single Virtual Patient** e.g. T2D with CHD
  - Generate diversity

- **Virtual Patient Cohort**
  - Prevalence weight

- **Virtual Population (VPop)**
  - Quantitative statistically-based predictions of clinical efficacy and candidate biomarkers

<table>
<thead>
<tr>
<th>Target Evaluation &amp; MOA</th>
<th>Biomarkers</th>
<th>Trial Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial evaluation of effects in a prototypical patient with average responses to approved therapeutics</td>
<td>Efficacy evaluation, Sensitivity analysis</td>
<td>N/A, N/A</td>
</tr>
<tr>
<td>Qualitative evaluation in a group of VPs across the range of mechanistic and phenotypic diversity</td>
<td>Efficacy evaluation, Sensitivity analysis</td>
<td>Define ranges of candidate biomarkers/panels</td>
</tr>
<tr>
<td>Quantitative statistically-based predictions of clinical efficacy and candidate biomarkers</td>
<td>Efficacy evaluation, Sensitivity analysis, Competitive differentiation, Combination with other Tx</td>
<td>Type 0, Type 1, Type 2, &quot;Break biomarker&quot; analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Protocol optimization, Qualitative identification of optimal target population</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Protocol optimization, Quantitative identification of optimal target population</td>
</tr>
</tbody>
</table>
CETP inhibitors have apparently beneficial effects on cholesterol but may not reduce CV risk

- The first CETP inhibitor, torcetrapib, was terminated in Phase III due to increased CV events believed to be a result of off-target increases in blood pressure
- Subsequently two new CETP inhibitors have entered late stage clinical trials to evaluate CV events
- The “class v molecule” hypotheses can be tested \textit{in silico} by representation of CETP activity at the mechanistic level as follows:
  - Create a VPop with baseline lipoprotein and plaque measures matching reported data for the terminated torcetrapib trial
  - Implement reported PK/PD for two new CETP inhibitors
  - Simulate theoretical clinical trials comparing statin alone with statin + new CETP inhibitors
  - Predict the relative risk within the VPop in response to the addition of each CETP inhibitor compared with statin alone
VPop matching baseline data captures the diversity seen in lipid profiles across various CETPi trials

VPop captures appropriate response to each simulated CETP inhibitor for measured lipoproteins
No overall CV risk benefit is seen in the simulated VPop in response to new CETP inhibitors

- Some VPs exhibit relative improvement with the addition of CETP inhibitor to the statin background, others are worse off
- Interrogation of the model results could determine and differentiate the mechanisms and biomarkers in responders and non-responders

Summary: CTS can be used to predict outcomes and define responders and non-responders

- Despite epidemiological data suggesting that HDL-C levels are associated with low CHD risk raising the level of HDL-C by inhibition of intravascular transfer between particles due to CETP activity in silico is not predicted to be beneficial in the patient population chosen for the Phase III clinical trials
  - Clinical Trial Simulations of Dyslipidemic Patients in a Mechanistic Model of Cardiovascular Disease Predict Little Impact on CHD Events by CETP Inhibitors. Wahba K et al., Circulation. 2011;124:A9560

- Recently a second CETP inhibitor was withdrawn from late stage development, due to a “lack of clinically meaningful efficacy”
  - Six clinical trials were affected: dal-OUTCOMES, dal-OUTCOMES 2, dal-PLAQUE 2, dal-ACUTE, dal-PLAQUE (completed) and dal-VESSEL (completed)
  - The estimated cost of the six program may be upwards of $420M (estimate based on “Clinical Trials of Drugs and Biopharmaceuticals” by Lee et al)

- CTS analysis prior to trial design could be used to identify and enrich for responders and against non-responding sub-populations based on biomarkers measured at baseline or during an initial screening period
Case Study: Safety and efficacy of extended dosing for an EPO stimulating agent (Phase IV)

<table>
<thead>
<tr>
<th>Background</th>
<th>Need to demonstrate safety and efficacy of novel extended dosing schedule in a specific population of anemic patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
<td>Adaptive trial design has to be simple and effective while meeting safety and efficacy endpoints</td>
</tr>
<tr>
<td>Questions</td>
<td>What are the effects of the various trial design choices on efficacy and safety end-points and how can the trial design be optimized to achieve both efficacy and safety?</td>
</tr>
</tbody>
</table>
Example protocol design for trial and types of features necessary for long-term reduced dosing

Total trial duration = 40 weeks

- Escalate frequency if Hb drops below 10 g/dl
- Reduce frequency if Hb exceeds 11 g/dl
- Suspend dosing if Hb > 13 mg/dL

Simulated results of final protocol design

- CTS suggests it is possible to dose less frequently and maintain patient safety
- Monthly dosing may not be as efficacious as more frequent doses, however
- Monthly dosing resulted in less dramatic Hb fluctuations

<table>
<thead>
<tr>
<th>Weeks on trial</th>
<th>Dose (U)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protocol</th>
<th>QW</th>
<th>Q2W</th>
<th>Q4W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of subjects whose mean Hb over last 4 weeks exceeds 10.5 g/dL</td>
<td>90-%</td>
<td>90-%</td>
<td>75-%</td>
</tr>
<tr>
<td>Maximum Hb level (g/dl)</td>
<td>11.2 ± 0.9</td>
<td>11.3 ± 0.8</td>
<td>10.9 ± 0.7</td>
</tr>
<tr>
<td>Proportion of subjects whose rate of Hb rise in 2 weeks exceeds 1 g/dl</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Maximum rate of Hb rise over 2 weeks</td>
<td>0.4 ± 0.1</td>
<td>0.5 ± 0.1</td>
<td>0.6 ± 0.1</td>
</tr>
</tbody>
</table>
Summary: CTS rapidly optimizes potential clinical trial protocols for efficacy and safety

- Patients often need frequent visits to the doctor, monitoring and dose adjustments to maintain key biological endpoints (e.g. hemoglobin (Hb) in CKD or FBG/HbA1c in diabetes)

- Reduction of dosing frequency is desirable but necessitates longer clinical trials especially when the entrance criteria are broad (standard is to initiate therapy when Hb is between 7 and 11 g/dL) and each individual patient has a different rate of progression

- CTS simulations of alternate trials allow the evaluation of virtual patients with a full range of different entry Hb and different disease progression rates
  - Each set of trial simulations takes 2-3 hours and only a day or two to analyze
  - Both safety and efficacy can be determined

Summary

- CTS simulations enable the mathematical representation of large scale biological processes, complex diseases and whole body systems to predict individual therapeutic responses and better understand mechanisms for sub-population stratification
  - In contrast to PK/PD approaches, which make the simplifying assumption that the body can be represented as a series of 3-4 compartments, a mechanistic model allows a much more detailed understanding of 1) the mechanism of action and 2) the extent and impact of individual patient variability
  - Established mechanisms fail 35% less often than poorly understood mechanisms (McKinsey perspectives on drug and device R&D 2012 “Escaping the sword of Damocles: Toward a new future for pharmaceutical R&D”)
  - Dr. Steve Paul, former head of Lilly Research Laboratories, quantified the value of a 1% reduction in phase II attrition to be $25M for a company like Lilly

- CTS can improve therapeutic development by
  - Early identification of therapeutics which are not significantly better than current standards of care
  - Identifying patients/sub-populations most likely to benefit
  - Improving trial design and duration in terms of dosing frequency and length etc.
Opportunities and Challenges: A Reality Check

Christy Chuang-Stein
Statistics, Pfizer Inc

The Fourth Hurdle

There was general agreement that the fourth hurdle was the one to look out for.

Source: Chrissie Fletcher
We Wish It Were That Simple

- How sure are we that the defined subgroup will benefit to a greater extent, or the remaining subgroup will not benefit at all or derive minimum benefit?
- Is the subgroup identified through an exploratory analysis of a study demonstrating no overall treatment effect?
- If a subgroup is defined by a molecular target, do we have a reliable test with acceptable specificity and sensitivity to identify the subgroup?
- Often, the real danger is we think we know the answer before we have the data; or the data we rely on are flawed.

Source: Kevin Carroll (2011), modified.
The Story of Tarenflurbil

- Tarenflurbil is a selective Aβ42-lowering agent that has been shown in vitro and in vivo to reduce Aβ42 production in favour of less toxic forms of Aβ.
- A Phase 2 trial (12 month DB treatment with 12-month extension) enrolled 210 patients between Nov 3 2003 and April 24 2006 in 31 sites in Canada and UK. The study included two doses (400 vs 800 mg) and placebo.
- A phase 3 study (800 mg vs placebo) began on Feb 21 2005 at 133 sites in US. The study initially enrolled both mild and moderate AD patients. After analysis of phase 2 data indicated that patients with mild AD responded better, enrolment was restricted to mild patients only.
- Efficacy results in 1684 mild AD patients were flat.


False Positive

- Multiplicity
  - With multiple subgroup analyses, probability of a false positive finding substantial.
  - With 10 independent tests ($\alpha=0.05$), chance of at least one false positive > 40%.

Simulation Study of Country Differences

- Hypothetical study
  - 4000 patients in 20 countries (200 patients each) with a control arm risk of 20% and an experimental arm risk of 15%
  - Homogenous absolute risk reduction of 5% in all countries.

Source: Marschner (DIA Annual Meeting).

Randomize-all Design

- Pros:
  - Able to estimate the benefit and risk of experimental trt vs. control in each subgroup.

- Cons:
  - The subgroups of interest need to be similarly prevalent, or the targeted subgroup needs to be more prevalent
  - For molecular target, biomarker status for some patients may be missing due to sample or assay problems
Some Adaptations for All-comers Design

- **Adaptive Selection Design**
  - Enroll all comers to start with; select a subgroup based on an interim analysis and limit future enrollment to the subgroup.

- **Adaptive Signature Design**
  - Enroll all comers; use 0.04 to test for all patients or use the first half to search for the subgroup most likely to benefit and test this subgroup using the second half at the 0.01 level (Freidlin and Simon, 2005).

- **Biomarker Adaptive Threshold Design**
  - Assay is ready for use except for the cutoff (Jiang, 2007).

Current State of Adaptive Designs

- Both FDA and EMA have a guidance on adaptive designs.
- There is much more flexibility for adaptations in the learn (exploratory) phase where the need to strongly control the false positive rate at the study level is much less.
- Common adaptive designs such as group sequential designs and blinded sample size re-estimation based on nuisance parameters are well accepted for phase 3 trials.
- Adaptive designs require a lot of upfront planning including extensive modeling and simulation (M&S) work. M&S allows us to critically examine many aspects of clinical trials (e.g. inclusion/exclusion criteria, dropout, enrichment).
- Adaptive designs should not be viewed as a means to salvage poorly-designed studies.
Success with M&S at Pfizer: Story 1

- Developing a pain medication
  - SC-75416 was a third-generation COX-2 inhibitor. The first POC trial failed to demonstrate a desirable advantage over ibuprofen using total pain relief at 6 hours (TOTPAR6) in the dental pain model.
  - Analysis showed that the blood level of SC-75416 was under the target level.
  - Conducted PK-PD modeling based on data from 1st and 2nd generation of COX-2 inhibitors and the POC study to estimate effect if the target blood level could be reached.
  - Developed a new formulation and conducted a new study. The observed benefit in the new study was consistent with the prediction.


Success with M&S at Pfizer: Story 2

- Developing an anti-coagulant
  - Used data on in vitro biomarker (inhibition of thrombin generation TG), clinical efficacy (venous thromboembolic event, VTE) and side effect (major bleeding, MB) in 43 trials and > 24,000 patients to build a model relating potency for VTE (MB) as a function of that for the inhibition of TG.
  - For a desirable VTE rate and acceptable MB rate, estimated the likely dosing range for a new product candidate based on the “same-mechanism” model and “all-mechanism” model.
  - Designed an adaptive design and investigated the dose range based on same-mechanism model. The study could add higher doses in the range estimated by the all-mechanism model, if necessary.
Success with M&S: Story 2 (cont)

**How the story ended**

- Patients were initially randomized to five doses and an active comparator.
- Three interim analyses were conducted. The lowest dose (among remaining doses) was dropped after each interim analysis.
- Two pre-specified higher doses, still within the predicted range under the “all-mechanism” model, were added.
- The entire 95% confidence interval for the dose with the target efficacy and safety is within the range predicted by the “same-mechanism” model.

EMA-EFPIA Workshop on M&S

- The workshop, held on Nov 31 – Dec 1 2011 in London, was to discuss the role and scope of M&S in drug-development both from the developer’s and the regulator's perspectives.
- The workshop offered some good case studies on how M&S was used to support decisions. For example, regulatory approval of doses that were not tested in clinical trials.

Opportunities

- We have clearly seen successes where the biology is well understood, a critical relationship is repeatedly demonstrated or where predictive animal models exist.
- We have come a long way in the technical implementation of M&S. We should certainly consider applying them to economic evaluations.
- Our phase 3 studies tend to get bigger. They offer a great opportunity to collect healthcare usage data for economic assessment.
- Cost-effective assessment is strongly related to efficacy assessment. Conducting economic assessment for subgroups is a natural next step. It supports reimbursement decisions.

Challenges

- Need data, data and data. Need to understand the strength and weaknesses of the data.
- Need to understand assumptions behind the models and the validity of the assumptions at the individual patient level.
- Need to properly quantify uncertainties and acknowledge what we don’t yet know.
- Need biological/clinical rationale for the models employed and subgroups identified. More importantly, the identified model should be applicable to similar programs.
Final Words

- Pharmaceutical sponsors have always embraced scientific approaches that will make product development more effective and product commercialization more successful.

- Working in the innovative pharmaceutical industry requires a heavy dose of optimism and enthusiasm. We need to temper them with sound reasoning and reality check.

Q&A

Ananth Kadambi and Bob Leipold will be available at the UBC booth (#77-80) following the workshop for additional discussion.