Outline: Simulated Treatment Comparison
- Description of method
- Major analytical steps
- Illustration of the method
- General guidance on application of STC

Typical Situation
- Pivotal (index) Phase III study have shown that treatment A is efficacious compared to placebo
- Preparing for reimbursement and access submissions
- A competing treatment (B) is already on the market
- Agencies require comparative evidence for A vs. B
- No head-to-head study has been done

Standard Approach
- Mixed treatment comparison (MTC) or network meta-analysis
  - Gather all studies that include evidence about the efficacy of A and B that could be "linked" through common comparators
  - Extract relative effect estimates of A vs. plc and B vs. plc
  - Perform the MTC to derive summary estimate of A vs. B
- Advantage
  - All available evidence is incorporated in the analysis
  - Uncertainty in study-specific estimates properly taken into account
- Challenge
  - Handling heterogeneity
  - Applicability of findings
Heterogeneity in MTCs/Meta-analyses
- Are all of the studies measuring the same effect?
- Dealing with heterogeneity:
  - Weed out studies that are obviously different, and pool the rest
  - Test for heterogeneity, and if detected, incorporate a random-effect in the estimation of the summary effect
  - Assume variability is completely random
- The question being answered: What is the average difference between A and B?
- Does the average difference apply to the index trial?
- What would the difference have been if B was included in the index trial?

Simulated Treatment Comparison
General Idea
- Estimate the expected difference between A and B, had the index study included an arm that was randomized to B
  - Predict key outcomes with treatment B in the index trial (so maintaining population and setting)
  - Derive estimates of comparative metrics (e.g., hazard ratios, odds ratios, mean differences)

Simulated Treatment Comparison
Main Steps
1. Derive predictive equations for the key outcomes of the index trial of A
2. Calibrate index equations to predict outcomes for B
   - Identify a study of B that is comparable to the index study of A
   - Use data on outcomes with B (ideally patient-level, but more likely published information) to calibrate equations to predict outcomes for B
   - Taking into account differences between the two populations
3. Build a simulation of the index trial based on calibrated equations
   - Generate predictions of outcomes with both treatments
   - Calculate effect measures or other metrics of interest

Example
- Index trial compared new treatment (A) to standard care
  - Key outcome: Hospitalization
  - Want to compare A to leading competitor B
  - Head-to-head study not yet available
  - Additional challenge: studies of B have not looked at hospitalization
  - BUT, have access to patient level data of an important study of B where hospitalizations were tracked but not fully studied

Caro & Ishak (2010), Pharmacoeconomics.
Step 1: Derivation of Predictive Equations

- **Parametric Survival Analysis**
  - Relates risk to determinants
  - Allows prediction of specific event times

- **Process**
  - Identify best fitting parametric distribution for the data
    - Test exponential, Weibull, Gompertz, log-normal, log-logistic
  - Identify predictors of risk (among those shown above)
  - Univariate analysis testing association of each predictor alone
  - Multivariate model with significant predictors from previous step
    - Trimmed final model including only significant or clinically important predictors in the multivariate model
  - Check model fit against observed data

---

**HR for Trt A vs. Placebo** replicated observed result to two decimals

**Statistical Criteria for Assessing Model Fit**

<table>
<thead>
<tr>
<th>Model</th>
<th>Null</th>
<th>Cohort</th>
<th>df</th>
<th>PDC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>-1221.85</td>
<td>-1220.99</td>
<td>3</td>
<td>2445.71</td>
<td>2458.591</td>
</tr>
<tr>
<td>Weibull</td>
<td>-1220.99</td>
<td>-1219.63</td>
<td>3</td>
<td>2446.328</td>
<td>2465.704</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-1220.84</td>
<td>-1219.92</td>
<td>3</td>
<td>2445.836</td>
<td>2465.156</td>
</tr>
<tr>
<td>Gamma</td>
<td>-1220.9</td>
<td>-1219.99</td>
<td>4</td>
<td>2447.985</td>
<td>2473.745</td>
</tr>
<tr>
<td>Log-Normal</td>
<td>-1223.62</td>
<td>-1222.39</td>
<td>3</td>
<td>2450.781</td>
<td>2470.101</td>
</tr>
<tr>
<td>Log-logistic</td>
<td>-1221.07</td>
<td>-1220.13</td>
<td>3</td>
<td>2446.257</td>
<td>2465.576</td>
</tr>
</tbody>
</table>

**Step 1. Predictive Equation for Time to Hospitalization**

\[
\lambda(T_{rt}, X) = \alpha + \beta T_{rt} + \theta X
\]
Log-Normal Risk Equation for Hospitalization

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
<th>Predictors</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A vs. Plc</td>
<td>0.502</td>
<td>0.085</td>
<td>0.000</td>
<td>X1</td>
<td>-0.329</td>
<td>0.124</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>-0.357</td>
<td>0.12</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>-0.36</td>
<td>0.171</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>-0.235</td>
<td>0.082</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>0.231</td>
<td>0.133</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>0.007</td>
<td>0.064</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>-0.206</td>
<td>0.123</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>-0.013</td>
<td>0.006</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>0.507</td>
<td>0.094</td>
<td>0.050</td>
</tr>
</tbody>
</table>

Note: positive coefficients are indicators of a protective effect
Estimate of the intercept and scale parameter not shown.

Step 2. Calibrating Equation to Predict for B

$\lambda(t) = e^{\alpha + \beta T_{Trt} + \theta X}$

Where:
- $\lambda(t)$: hazard rate at time $t$
- $T_{Trt}$: treatment indicator
- $X$: covariates

Step 3: Calculate SSE between target and predicted curves

Target: How patients in study A would do if they received trt B

(\(\delta = 0\))

Placebo

(\(\delta = 0\))

(\(\delta = 0.3\))
Comparative Results from Equations

<table>
<thead>
<tr>
<th>Ratio Mean Hosp Times</th>
<th>Mean Hosp Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>A vs. P: 1.65 (Original Equation)</td>
<td>1.06 (Calibration)</td>
</tr>
<tr>
<td>A vs. B: 1.56 (Calibration)</td>
<td></td>
</tr>
</tbody>
</table>

Step 3. Simulation of Index Trial with Added Arm for B

Why Simulate?
- Can replicate original trial, but also modify certain aspects of the study (e.g., patient profile)
- Used to generate transition probabilities to inform economic modeling
- Used as a tool to replicate sampling and generate variability bands around predicted outcome
  - These are not confidence bands; they do not reflect all sources of uncertainty. Only sampling variability in the conduct of the trial with three arms.
- Trial simulation framework to help with design of future studies
  - E.g., STC after Phase II to help plan phase III

Simulation of the Index Trial

Create patients → Read in all patients from Index trial → Assign Trt A → Use calibrated equations to derive event times (hospitalization, and death) → Assign Trt B → Determine next event & delay until next event → Process event, record time → Death? → Record Disease-related Death → Exit → Record All-Cause Death → Model End?

Example of Trial Simulation Results

<table>
<thead>
<tr>
<th></th>
<th>A from Trial</th>
<th>A Model</th>
<th>P Trial</th>
<th>P Model</th>
<th>B Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosp</td>
<td>29.3%</td>
<td>26.4%</td>
<td>36.9%</td>
<td>30.8%</td>
<td>34.9%</td>
</tr>
<tr>
<td>Death from any cause</td>
<td>5.0%</td>
<td>5.0%</td>
<td>6.0%</td>
<td>5.6%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Death due to Disease</td>
<td>2.7%</td>
<td>2.6%</td>
<td>3.9%</td>
<td>3.5%</td>
<td>3.3%</td>
</tr>
</tbody>
</table>
What if we didn’t have patient-level data for B?

**Scenario 1:** have KM curve for outcome for Trt B

1. Obtain an appropriate reference point (percentile) for B
2. Adjust placebo curve to characteristics of patients in the study of Trt B
3. Calibrate equation to cross the observed survival point in the study of Trt B

Adjusts for differences in characteristics between studies

**Scenario 2:** Only have a single point from KM curve

1. Step 3: Calibrate equation to cross the observed survival point in the study of Trt B

Finding a Comparable Study for Trt B

- What do we mean by “comparable”? How similar do the studies have to be? What differences are important and which are not?
- Differences in baseline characteristics not a limiting factor. Can handle analytically, but can’t rule out confounding from unmeasured characteristics
- Problematic differences
  - Look for entire subsets are not systematically excluded in one of the studies
  - No extreme difference in FU duration (e.g., 5yr FU in index vs. 6mo FU)
  - Studies with very different treatment protocols
  - Different measurements techniques
Choosing from Multiple Studies for Comparator

- What if there are multiple studies for the comparator?
  - Choose one? Combine? Repeat with each study?
- Choose one if
  - Differences between studies point to a closer match to the index trial
- If no clearly better choice, combining may be a good option
  - Incorporate all pertinent evidence
  - Combining published data may force some simplification: choosing a specific point on survival curves as a reference value for calibration (as opposed to using the entire curve)
- Or repeat analyses with each study
  - To test robustness of results

Using the Reference Arm of Comparator Studies

- The reference arm of the comparator study can be very useful if the treatment assigned in this group is the same as the reference group in the index trial
- Outcomes in the reference group of the two studies provide a means of checking the comparability of studies
  - Similar outcomes in the two reference arms (after adjusting for population differences) implies comparability of studies
  - Differences (after adjusting for population differences) implies “study effects” that may distort comparative results
    - May be due to calendar effects
    - Different outcome definitions
    - Etc.
  - Can adjust for the study effect in STC result using difference between reference curves

Uncertainty of the STC Result

- Variability bands shown earlier only reflect sampling variation, not uncertainty
- Uncertainty in the results stems from
  - Uncertainty in the predictions from the predictive equations
  - Uncertainty in the comparator results used for calibration
- Quantifying the uncertainty of the STC result requires propagating the uncertainty in predictions and comparator results through the estimation process
  - The simulation framework allows for this be done very easily
  - This would produce a distribution of results based on which a standard error and confidence interval can be derived
STC with Other Types of Outcomes

- The example illustrates use of STC with time-to-event outcomes
- The same can be done with continuous (e.g., weight change) or dichotomous (e.g., controlled BP)
- Strategy would be the same, with similar considerations
  - Comparability of studies
  - Criteria for calibration: specific measure (e.g., mean vs. entire curve)
  - Uncertainty of estimates

Summary

- STC is an alternative approach to deriving indirect comparison
  - Won’t always be feasible or appropriate
  - Careful review and comparison of available data sources is an important first step
- STC answers a different question
  - Provides a comparison within the context of a specific study
  - Can be complementary to MTC
  - Differences in results does not imply one or the other is incorrect
- Work is ongoing to develop and refine the approach through applications