MAKING SENSE OF NOVEL APPROACHES FOR INDIRECT COMPARISON: SIMILARITIES AND DIFFERENCES OF SIMULATION AND MATCHING BASED APPROACHES

K. Jack Ishak PhD, Hemant Phatak PhD, Cristina Masseria, PhD

ISPOR 18th Annual European Congress

November 9th, 2015

Acknowledgments

▪ Acknowledgments
We’d like to recognize the very important contributions made to this work by our colleagues Michael Rael PhD, Tereza Lanitis MSc, and Meredith Hoog MPH. They were involved in the conceptualization, execution and interpretation of the analyses.

▪ Conflicts of Interest:
J Ishak, M Rael, T Lanitis and M Hoog are employees of Evidera which received funding from Pfizer/BMS in connection with conducting this study and with the development of this workshop. H Phatak was an employee of BMS at the time of the study; he owns BMS stocks. C Masseria is an employee of Pfizer with ownership in stocks.
Learning Objectives

- To assist researchers and decision makers in gaining a better understanding of simulated treatment comparison (STC) and matching-adjusted indirect comparison (MAIC) as alternative and complimentary approaches to traditional network meta-analysis (NMA).

- Illustrate applications of these methods with an example based on treatments for non-valvular atrial fibrillation (NVAF) where all three approaches have been applied.

- Understand similarities and differences between the methods and their underlying assumptions

Outline

- Case Study: Comparison of treatment A and treatment B for preventing bleeds and stroke

- Overview of MAIC and STC and how they were applied in the case study

- Results from analyses of the case study

- Generalizing beyond the example

- Questions & Discussion
Context: Treatment of NVAF

- Anticoagulation treatment is recommended to mitigate the risk of stroke in patients with NVAF.
  - Until recently mitigation efforts focused on treatment with vitamin K antagonists (VKA; such as warfarin) or antiplatelet therapy (such as aspirin).
  - VKA treatment is limited by a monitoring requirement, food-drug and drug-drug interactions, and bleeding effects.

- Non-VKA oral anticoagulants (NOACs) exhibit equivalent or superior efficacy and/or safety to VKA without the monitoring requirement
  - Recommended by European Society of Cardiology guidelines and the American Stroke/Heart Association guidelines
  - No clear recommendations for the use of one NOAC versus another

- In the absence of head-to-head trials, several indirect treatment comparisons have examined relative effectiveness using network meta-analysis (NMA) techniques
  - The results of these indirect comparisons vary slightly depending on outcome measures and approach used.
Application to Indirect Comparisons of NOACs

- Discussion in this workshop will focus on two specific NOACs, which we will refer to as treatments A and B
- Treatments A and B have shown positive results in clinical trials.
  - Treatment A was shown superiority to dose-adjusted VKAs regarding reduction of stroke and systemic embolism (SSE) and major bleeding.
  - Treatment B was non-inferior to VKA in the prevention of SSE, with no significant difference between the treatments in the risk of major bleeding.
- Analyses relied on phase III registrational trials of the two treatments
  - Patient level data available from trial of treatment A
  - Published data available for trial of treatment B

Comparison of the Trials

- The trials of A and B were randomized and double-blind studies with follow-up duration of around two years
- The study populations were similar in age and gender distribution, but differed in their severity
- Outcomes in the W arms also differed
- Definition of bleeds were not the same in the two studies
Potential Limitations of Published NMAs

- NMAs may be impacted by
  - Differences in patient baseline characteristics
  - Performance of VKA arms in terms of time in therapeutic range.

- These differences can cause systematic variation between outcomes observed in studies
  - Better outcomes in the VKA arms can lead to an underestimation of the effect of the NOAC
  - Healthier (or sicker patients may derive) more or less benefit from treatments, causing systematic differences in effects estimated in different trials

- NMAs use random effects to account for the impact of these differences
  - Treats these as source of additional random variation.

Targeted Comparisons with MAIC and STC

- Objective: Explicitly account for differences in patient populations and reference arms in indirect comparisons of NOACs using matched adjusted indirect comparison (MAIC) and simulated treatment comparison (STC)
  - These methods compare outcomes directly between specific treatment arms of interest (e.g., NOAC arms of different trials), adjusting for differences in baseline characteristics of the trials
  - They can also adjust for other differences between the trials when these have common comparator arms (e.g., VKA arms of NOAC trials).

- MAIC and STC use the same data elements to derive adjusted comparisons
  - They differ in the analytical process used to adjust for differences in baseline characteristics.
Compatibility of the Trials for Comparison

- Trials deemed compatible for comparison due to similarity of designs and availability of data to adjust for differences in populations

- Populations differed in risk profile – Treatment B study had more severe profile
  - MAIC and STC designed to explicitly account for this

- Patients randomized to treatment W in study for treatment A study were in INR range for longer amount of time
  - May reflect other background differences between the trials
  - Since reference arms are the same, we can quantify and adjust for such differences.

Overview of MAIC and STC

Jack Ishak, PhD
Simplified Example Based on Trial Data

<table>
<thead>
<tr>
<th>Trial A</th>
<th>Treatment A</th>
<th>Treatment W(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Rate</td>
<td>2.1 (1.91, 2.37)</td>
<td>3.1 (2.82, 3.39)</td>
</tr>
<tr>
<td>% Severe</td>
<td>30.2%</td>
<td>30.2%</td>
</tr>
<tr>
<td>Rate Ratio</td>
<td>0.69 (0.60–0.80)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trial B</th>
<th>Treatment B</th>
<th>Treatment W(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Rate</td>
<td>3.6 (3.26, 3.97)</td>
<td>3.4 (3.08, 3.76)</td>
</tr>
<tr>
<td>% Severe</td>
<td>87.0%</td>
<td>86.9%</td>
</tr>
<tr>
<td>Rate Ratio</td>
<td>1.04 (0.90–1.20)</td>
<td></td>
</tr>
</tbody>
</table>

Simple “NMA” of the Trials

\[
RR_{AW} = \frac{Y_A}{Y_{W(A)}}
\]

\[
RR_{BW} = \frac{Y_B}{Y_{W(B)}}
\]

\[
RR_{AB} = \frac{RR_{AW}}{RR_{BW}} = \frac{Y_A/Y_{W(A)}}{Y_B/Y_{W(B)}} = \frac{0.69}{1.04} = 0.66
\]

Adjusts comparison for differences between studies
Non-Common Reference Arms

\[
RR_{AW} = \frac{Y_A}{Y_{W(A)}}
\]

\[
RR_{BW} = \frac{Y_B}{Y_{W(B)}}
\]

\[
RR_{BC} = \frac{Y_B}{Y_C}
\]

\[
RR_{AB} = \frac{RR_{AW}}{RR_{BC}} = \frac{Y_A/Y_{W(A)}}{Y_B/Y_C} = \frac{Y_A/Y_B}{Y_{W(A)}/Y_C}
\]

Reflects differences between studies as well as difference between C and PBO

Targeted Comparison of A vs. B

\[
RR_{AW} = \frac{Y_A}{Y_{W(A)}}
\]

\[
RR_{BC} = \frac{Y_B}{Y_C}
\]

\[
RR_{AB} = \frac{Y_A}{Y_B}
\]

Confounded by differences in populations
Targeted comparison: Use data from trial arms of interest, with adjustment for confounding.
# Targeted Comparison of A vs. B

<table>
<thead>
<tr>
<th>Trial A</th>
<th>Treatment A</th>
<th>Treatment W(A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Rate</td>
<td>2.1 (1.91, 2.37)</td>
<td>3.1 (2.82, 3.39)</td>
</tr>
<tr>
<td>% Severe</td>
<td>30.2%</td>
<td>30.2%</td>
</tr>
<tr>
<td>Rate Ratio</td>
<td>0.69 (0.60–0.80)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trial B</th>
<th>Treatment B</th>
<th>Treatment W(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Rate</td>
<td>3.6 (3.26, 3.97)</td>
<td>3.4 (3.08, 3.76)</td>
</tr>
<tr>
<td>% Severe</td>
<td>87.0%</td>
<td>86.9%</td>
</tr>
<tr>
<td>Rate Ratio</td>
<td>1.04 (0.90–1.20)</td>
<td></td>
</tr>
</tbody>
</table>

\[
RR_{AB} = \frac{Y_A}{Y_B} = \frac{Y_A(X_A)}{Y_B(X_B)} \quad \text{if} \quad X_A \neq X_B \quad \text{so the comparison is biased.}
\]

\[
RR_{AB} = \frac{Y_A}{Y_B} = \frac{Y_A(X_A)}{Y_B(X_B)}
\]

---

## Methods for Adjusting for Potential Confounding

- **Two approaches**
  - Matching Adjusted Indirect Comparison (MAIC)
  - Simulated Treatment Comparison (STC)

- **MAIC** adjusts for differences between populations by weighting patients in the index trial so that the average characteristics match the comparator’s population
  - Adjusted outcomes reflect what would be expected with treatment A in the comparator’s population, and thus, can be used to calculate a relative effect estimate
  - Adjusted outcomes can be compared directly with treatment B to derive a relative effect estimate

- **STC** generates the adjusted outcomes with treatment A using a predictive equation
  - Predicted values are generated for the comparator’s trial population
  - Adjusted outcomes reflect what would be expected with treatment A in the comparator’s population, and thus, can be used to calculate a relative effect estimate
Illustration of MAIC

- **Data from trials A and B**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe</td>
<td>30.2%</td>
<td>87.0%</td>
</tr>
<tr>
<td>Rate (Y)</td>
<td>2.1</td>
<td>3.6</td>
</tr>
</tbody>
</table>

- **The aim is to reweight patients in trial A, so that the percent severe in the weighted population is 0.87**

  \[
  Y_A(X_B) = \frac{\sum_{i=1}^{N} w_i Y_{Ai}}{\sum_{i=1}^{N} w_i} = 2.7
  \]

  \[
  RR_{AB} = \frac{Y_A(X_B)}{Y_B(X_B)} = \frac{2.7}{3.6} = 0.75
  \]

Generalization to Multiple Matching Variables

- **Must determine weights that will simultaneously match multiple variables**
  - Propensity score-type (logistic regression) equation is derived to predict probability of patient coming from the index vs. comparator trial
  - Using patient level data from the index trial and mean characteristics for comparator

- **Weights derived from the predicted odds from the propensity equation**
  - When index and comparator populations are very similar, weights will have a fairly uniform distribution in the index population
  - When differences are large, weights may be concentrated in subsets of patients that are ultimately driving the adjusted result

- **The extent of differences in populations determines information loss due to matching**
  - “Effective sample size” is a measure derived from the weights to reflect the fraction of original sample contributing to adjusted outcome.
Illustration of STC

- STC uses predictive equations to derive adjusted rates
- Develop a Poisson equation to relate severity to outcome in trial A
  \[ \log(Y_A(X)) = \beta_0 + \beta_1 \text{Severe} \]
- We obtain:
  \[ \log(Y_A(X)) = 0.61 + 0.45 \times \text{Severe} \]
- Then:
  - Predicted Rate in severe: \( \exp(0.61+0.45) = 2.9 \)
  - Predicted Rate in non-Severe: \( \exp(0.61) = 1.8 \)

  \[ Y_A(X_B) = 0.87 \times 2.9 + 0.13 \times 1.8 = 2.7 \text{ (per 100 pyrs)} \]
- Adjusted rate ratio: \( RR_{AB} = \frac{Y_A(X_B)}{Y_B(X_B)} = \frac{2.7}{3.6} = 0.75 \)

Generalization to Multiple Predictor Variables

- In typical situations, equations will have to include multiple predictors
  - Multivariate equation, properly tested for fit
- With non-linear outcomes like rates, probabilities or time-to-event outcomes, the adjusted outcome should not be calculated by plugging in the mean values for the comparator
- Suppose the equation is: \( \log(Y_A(X)) = \beta X \)
  - We seek the adjusted rate:
    \[ \frac{1}{N} \sum Y_A(X_{B_i}) = \frac{1}{N} \sum \exp(\beta X_{B_i}) \]
    - This is the arithmetic mean rate with A in the comparator population.
  - Since we only have means of Xs, we can only calculate
    \[ Y_A(\bar{X}_B) = \exp(\beta \bar{X}_B) = \sqrt[N]{\prod \exp(\beta X_{B_i})} \]
    - This is the geometric mean rate with A in the comparator population.
- The arithmetic mean is always larger than the geometric mean.
Adjusted Outcomes from Simulated Patient Profiles

- To overcome the issue, patient profiles can be simulated to reflect the comparator population

- That is:
  - Simulate $X_{Bi}$ so that mean($X_{Bi}$) match the means in comparator population
  - But must ensure the combination of values assigned to Xs reflect natural correlations between risk factors

- Approach:
  - Simulate by taking draws from a multivariate normal distribution setting means to those from the comparator population and correlation matrix from the index trial
  - Challenge: dealing with categorical predictors like history of disease or use of treatment
    - Simulation is done on probability of having a given characteristic – i.e., the simulated value is a probability
    - Characteristic is assigned based on whether the simulated value exceeds the prevalence in the comparator population.

- Fundamental assumption:
  - Correlation between risk factors in index trial applies to comparator’s population
  - Same assumption is made in MAIC.

Study Effects

- We call study effects differences in outcomes between treatments A and B due to factors other than the effects of the treatment and characteristics of the population

- Other potential causes
  - Residual confounding due to variables available in only one or neither study
  - Differences in study setting (calendar periods, geographic locations)
  - Differences in outcome definitions or design features (e.g., blinding, etc.)

- Study effects can be quantified when studies have common comparators by applying STC/MAIC on the comparator arms
  - No study effects implies close compatibility
  - Large study effects may be indicative of incompatibility
  - Otherwise, the indirect comparison of A vs. B can be adjusted accordingly.
### Uncertainty for Comparison of W(A) vs. W(B)

<table>
<thead>
<tr>
<th></th>
<th>A vs. B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td></td>
</tr>
<tr>
<td>Rate in Trial A</td>
<td>2.1</td>
</tr>
<tr>
<td>(N~9K)</td>
<td></td>
</tr>
<tr>
<td>Rate in Trial B</td>
<td>3.6</td>
</tr>
<tr>
<td>(N~7K)</td>
<td></td>
</tr>
<tr>
<td>MAIC (N~3,500)</td>
<td></td>
</tr>
<tr>
<td>Adjusted Rate for</td>
<td>2.7</td>
</tr>
<tr>
<td>arm in Trial A</td>
<td></td>
</tr>
<tr>
<td>Rate Ratio</td>
<td><strong>0.75</strong></td>
</tr>
<tr>
<td>STC</td>
<td></td>
</tr>
<tr>
<td>Adjusted Rate for</td>
<td>2.7</td>
</tr>
<tr>
<td>arm in Trial A</td>
<td></td>
</tr>
<tr>
<td>Rate Ratio</td>
<td><strong>0.75</strong></td>
</tr>
</tbody>
</table>

Adjusting for Study Effects: \( \frac{RR_{AB}}{RR_{WA/WB}} = \frac{0.75}{1.16} = 0.65 \)

---

### Results from Analyses of A vs. B

_Hemant Phatak, PhD_
Baseline Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
<th>Treatment W in &quot;A&quot; Study</th>
<th>Treatment W in &quot;B&quot; Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>~9K</td>
<td>~7K</td>
<td>~9K</td>
<td>~7K</td>
</tr>
<tr>
<td>Male sex</td>
<td>0.646</td>
<td>0.603</td>
<td>0.650</td>
<td>0.603</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>0.246</td>
<td>0.188</td>
<td>0.245</td>
<td>0.188</td>
</tr>
<tr>
<td>Europe</td>
<td>0.403</td>
<td>0.532</td>
<td>0.404</td>
<td>0.533</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-75</td>
<td>0.388</td>
<td>0.331</td>
<td>0.387</td>
<td>0.334</td>
</tr>
<tr>
<td>&gt; 75</td>
<td>0.312</td>
<td>0.438</td>
<td>0.311</td>
<td>0.436</td>
</tr>
<tr>
<td>Race white</td>
<td>0.826</td>
<td>0.830</td>
<td>0.825</td>
<td>0.835</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>83.931</td>
<td>82.100</td>
<td>84.144</td>
<td>81.600</td>
</tr>
<tr>
<td>Persistent or permanent atrial fibrillation</td>
<td>0.849</td>
<td>0.825</td>
<td>0.844</td>
<td>0.822</td>
</tr>
<tr>
<td>No Prior use of vitamin K antagonist (i.e. VKA naive)</td>
<td>0.428</td>
<td>0.377</td>
<td>0.428</td>
<td>0.375</td>
</tr>
<tr>
<td>CHADS₂ score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHADS₂≥2</td>
<td>0.358</td>
<td>0.130</td>
<td>0.358</td>
<td>0.131</td>
</tr>
<tr>
<td>CHADS₂≥3</td>
<td>0.302</td>
<td>0.870</td>
<td>0.303</td>
<td>0.869</td>
</tr>
<tr>
<td>Renal function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal, &gt;80 ml/min</td>
<td>0.414</td>
<td>0.322</td>
<td>0.415</td>
<td>0.313</td>
</tr>
<tr>
<td>Mild impairment, &gt;50 to 80ml/min</td>
<td>0.420</td>
<td>0.211</td>
<td>0.417</td>
<td>0.207</td>
</tr>
<tr>
<td>Prior history of Stroke, TIA or systemic embolism</td>
<td>0.192</td>
<td>0.549</td>
<td>0.197</td>
<td>0.546</td>
</tr>
</tbody>
</table>

Note: Analyses also adjusted for other comorbidities and use of treatments at baseline.

Baseline Characteristics after Matching in MAIC

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
<th>Treatment W in &quot;A&quot; Trial</th>
<th>Treatment W in &quot;B&quot; Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>~9K</td>
<td>~1.5K*</td>
<td>~7K</td>
<td>~9K</td>
</tr>
<tr>
<td>Male sex</td>
<td>0.646</td>
<td>0.603</td>
<td>0.660</td>
<td>0.603</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>0.246</td>
<td>0.188</td>
<td>0.245</td>
<td>0.188</td>
</tr>
<tr>
<td>Europe</td>
<td>0.403</td>
<td>0.532</td>
<td>0.404</td>
<td>0.533</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-75</td>
<td>0.388</td>
<td>0.331</td>
<td>0.387</td>
<td>0.334</td>
</tr>
<tr>
<td>&gt; 75</td>
<td>0.312</td>
<td>0.438</td>
<td>0.311</td>
<td>0.436</td>
</tr>
<tr>
<td>Race white</td>
<td>0.826</td>
<td>0.830</td>
<td>0.825</td>
<td>0.835</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>83.931</td>
<td>82.100</td>
<td>84.144</td>
<td>81.600</td>
</tr>
<tr>
<td>Persistent or permanent atrial fibrillation</td>
<td>0.849</td>
<td>0.825</td>
<td>0.844</td>
<td>0.822</td>
</tr>
<tr>
<td>No Prior use of VKA</td>
<td>0.428</td>
<td>0.377</td>
<td>0.428</td>
<td>0.375</td>
</tr>
<tr>
<td>CHADS₂ score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHADS₂≥2</td>
<td>0.358</td>
<td>0.130</td>
<td>0.358</td>
<td>0.131</td>
</tr>
<tr>
<td>CHADS₂≥3</td>
<td>0.302</td>
<td>0.870</td>
<td>0.303</td>
<td>0.869</td>
</tr>
<tr>
<td>Renal function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal, &gt;80 ml/min</td>
<td>0.414</td>
<td>0.322</td>
<td>0.415</td>
<td>0.313</td>
</tr>
<tr>
<td>Mild impairment, &gt;50 to 80ml/min</td>
<td>0.420</td>
<td>0.211</td>
<td>0.417</td>
<td>0.207</td>
</tr>
<tr>
<td>Prior history of Stroke, TIA or systemic embolism</td>
<td>0.192</td>
<td>0.549</td>
<td>0.197</td>
<td>0.546</td>
</tr>
</tbody>
</table>

* Effective sample size
### MAIC – Results

<table>
<thead>
<tr>
<th></th>
<th>Stroke and Systemic Embolism</th>
<th>Major Bleeding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NOAC (A Vs. B)</td>
<td>Treatment W (W(A) vs. W(B))</td>
</tr>
<tr>
<td>Observed rates in “A” trial</td>
<td>1.27 (1.11, 1.453)</td>
<td>1.6 (1.418, 1.805)</td>
</tr>
<tr>
<td>Observed rates in “B” trial</td>
<td>2.1 (1.863, 2.367)</td>
<td>2.4 (2.146, 2.685)</td>
</tr>
<tr>
<td>Crude rate ratios</td>
<td>0.605 (0.505, 0.724)</td>
<td>0.667 (0.565, 0.786)</td>
</tr>
<tr>
<td>Adjusted rate after matching for treatment A</td>
<td>1.87 (1.42, 2.46)</td>
<td>2.50 (1.97, 3.18)</td>
</tr>
<tr>
<td>MAIC RR for A vs. B</td>
<td>0.886 (0.658, 1.199)</td>
<td>1.043 (0.802, 1.357)</td>
</tr>
<tr>
<td>MAIC RR for A vs. B adjusted for Study Effects</td>
<td>0.851 (0.571, 1.268)</td>
<td>1</td>
</tr>
</tbody>
</table>

### STC - Results

- The STC equation for SSE included: region, type of AF, prior VKA use, CCB use, BB Use, CHADS-2, Renal function, history of stroke/TIA/SSE.

- The STC equation for MB included: gender, region, age, prior use of ASA, ARBs, and diuretics, renal function, history of stroke/TIA/SSE, hypertension and diabetes.

<table>
<thead>
<tr>
<th></th>
<th>Stroke and Systemic Embolism</th>
<th>Major Bleeding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NOAC (A Vs. B)</td>
<td>Treatment W (W(A) vs. W(B))</td>
</tr>
<tr>
<td>Observed Rates in “A” trial</td>
<td>1.27 (1.11, 1.453)</td>
<td>1.6 (1.418, 1.805)</td>
</tr>
<tr>
<td>Observed Rates in “B” trial</td>
<td>2.1 (1.863, 2.367)</td>
<td>2.4 (2.146, 2.685)</td>
</tr>
<tr>
<td>Crude rate ratios</td>
<td>0.605 (0.505, 0.724)</td>
<td>0.667 (0.565, 0.786)</td>
</tr>
<tr>
<td>Predicted Rate for “A” using Virtual Profiles</td>
<td>2.049 (1.244, 3.375)</td>
<td>2.591 (1.621, 4.141)</td>
</tr>
<tr>
<td>STC RR for A vs. B</td>
<td>0.976 (0.584, 1.63)</td>
<td>1.079 (0.666, 1.748)</td>
</tr>
<tr>
<td>STC RR for A vs. B adjusted for Study Effects</td>
<td>0.904 (0.447, 1.828)</td>
<td>1</td>
</tr>
</tbody>
</table>
Comparison of STC and MAIC Rate Ratios

<table>
<thead>
<tr>
<th></th>
<th>Stroke and Systemic Embolism</th>
<th>Major Bleeding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NOAC (A Vs. B)</td>
<td>Treatment W (W(A) vs. W(B))</td>
</tr>
<tr>
<td>Crude Rate Ratio§</td>
<td>0.605 (0.505, 0.724)</td>
<td>0.667 (0.565, 0.786)</td>
</tr>
<tr>
<td>NMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate Ratio (Mitchell et al)</td>
<td>0.91 (0.71 – 1.16)</td>
<td></td>
</tr>
<tr>
<td>MAIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective Sample Size</td>
<td>~1.5K</td>
<td>~1.5K</td>
</tr>
<tr>
<td>Rate Ratio§</td>
<td>0.888 (0.658, 1.199)</td>
<td>1.043 (0.802, 1.357)</td>
</tr>
<tr>
<td>Study-Effect-Adjusted RR§</td>
<td>0.851 (0.571, 1.268)</td>
<td>1</td>
</tr>
<tr>
<td>STC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate Ratio§</td>
<td>0.976 (0.584, 1.63)</td>
<td>1.079 (0.666, 1.748)</td>
</tr>
<tr>
<td>Study-Effect-Adjusted RR§</td>
<td>0.904 (0.447, 1.828)</td>
<td>1</td>
</tr>
</tbody>
</table>

§ Rate ratios are displayed as A vs. B and W(A) vs. W(B)

Summary of Findings

- **MAIC and STC give generally consistent results**
- **They suggest comparable risk of SSE**
  - Point estimates suggest slightly lower risk with treatment A but not statistically significant
- **Risk of MB was found to be lower with treatment A**
  - Estimate was not statistically significant with STC, but reached significance in MAIC
  - Both methods showed large study effects for this outcome favoring treatment A
    - May reflect different definitions of bleeds which was broader in study A
- **Results are consistent with what's been observed in past NMAs**
General Considerations and Conclusion

Jack Ishak, PhD

Application to Comparison of Treatment A vs B

- **MAIC**
  - Balancing weights were derived for the active arms and control arms separately
    - All variables available in both sources were considered, but some were omitted due to difficulties in matching
  - Adjusted rates (per 100 person-years) derived for treatment A
    - By applying weights in calculation of the rate
  - Comparative results
    - Adjusted rate ratios were calculated for A vs. B and comparing two W arms
    - Study effects adjusted as follows:
      \[
      \hat{Y}_A(\hat{x}_A) \div \hat{Y}_W(\hat{x}_W) = \hat{Y}_A(\hat{x}_A) \div \hat{Y}_{refB}(\hat{x}_W)
      \]

- **STC**
  - Predictive equations derived for SSE and MB using Poisson regression
    - Tested all available variables, selected predictive ones through univariate multivariate selection process
  - Adjusted rates for ARISTOTLE arms predicted from equations
    - Virtual profiles simulated to match trial population characteristics for phase III study for treatment B
  - Comparative results
    - Derived as in MAIC
    - Adjusted for patient characteristics and study effects
STC vs. MAIC: Considerations in Choosing an Approach?

- STC and MAIC use the same data, and both adjust for differences between populations of the two studies
  - Reasonable to expect consistent results with both approaches
  - Useful to apply both to assess robustness of results

- STCs require one equation per outcome, while MAIC require one set of weights per comparator
  - STC more practical if interested in multiple comparators and few outcomes
  - MAIC more practical if interested in multiple outcomes and few comparators

- Type of outcome being compared: Generating predictions from equations of non-linear outcomes (e.g., dichotomous or time-to-event) must be handled carefully.

STC vs. MAIC: Considerations in Choosing an Approach?

- Extent of differences between the index and comparator populations
  - If index population is very different or has a narrower distribution of risk factors than the comparator, MAIC may produce very low ESS
  - Similarly, the STC equation may not be reliable to predict in the comparator population
  - Careful review (with clinical insight) is required to determine whether the ESS can be improved by omitting certain variables from the matching list, or whether projections outside of the range of the index data are plausible
    - The latter may sometimes be easier to defend than omitting potential confounders from the analyses

- STC may be more practical in situations where predictive equations are required, say for economic evaluations
  - The same equations can serve as the basis for both the economic model and the STC
  - Equations also provide transparency in terms of which risk factors may be confounding the comparison and which can be ignored.
Use of Targeted Comparisons with Agencies

- Targeted comparisons are being applied increasingly more
  - Both MAIC and STC have been included in NICE Technology Assessments
  - MAICs
    - Multiple Myeloma: bortezomib vs. cyclophosphamide, thalidomide and dexamethasone (TA311)
    - Hepatitis C: simeprevir plus peginterferon alfa and ribavirin vs. peginterferon alfa and ribavirin (TA331)
  - STC
    - Renal Cell Carcinoma: axitinib vs. best standard of care (TA333)

- NMA continues to remain the gold-standard expected by agencies
  - But growing recognition that alternative approaches can be used
  - E.g., in the submission for dasatanib vs. nilotinib and imatinib, the reviewers cited published MAIC conducted by competitor and asked the manufacturer to consider same analyses

Summary

- MAIC and STC are powerful techniques that can overcome potential issues in standard NMA
  - Used increasingly more as stand-alone analyses, as well as in context of economic evaluations

- STC with non-linear outcomes requires simulating patients to derive correct adjusted results for the index treatment

- Case study allowed applying and comparing standard and novel approaches to indirect comparison

- Consistent results were observed with all approaches
  - Not entirely surprising
  - MAIC and STC used the same data
  - NMA has a network of evidence but mostly relies on W as linking arm

- Stability of results is encouraging and supports credibility of approaches.
For more information...

Poster
SIMULATED TREATMENT COMPARISON OF TIME-TO-EVENT (AND OTHER NON-LINEAR) OUTCOMES (PRM 208)
Wednesday, November 11th, 8:45-13:45

Thank You!