Win Ratio: A New Approach to the Analysis of Composite Endpoints in Clinical Trials based on Clinical Priorities

Duolao Wang, Cono Ariti, Tim Collier & Stuart Pocock
London School of Hygiene & Tropical Medicine

Clinical Trials Methodology Conference- 4th and 5th October, Bristol.
Topics

- Composite endpoint and its problem
- Our new approach
- Results from three clinical trials
- Questions
Major RCT’s in CV disease use composite endpoints as the primary outcome to assess the treatments efficacy.

- Mixture of two or three types of clinical events (e.g., CV death, non-fatal MI, non-fatal stroke etc)
- Analysis focuses on time to the first event
  - Usually Cox model, KM plots, log-rank tests used for reporting treatment effects
- Implicitly treat all contributory endpoints as equal
- Typically only takes account of the first occurring endpoint
  - Non fatal events occurring earlier in follow-up tend to get a higher priority than later more serious events and deaths
Our New Approach

- Takes into account clinical priorities
  - CV deaths are considered more important than non-fatal events and get first priority
- Recognises that patients have widely different risk profiles
  - Risk-matched pairs are used in the analysis
- Method is easy to use
- Provides a novel estimate of the treatment difference with CI and P-value
Consider the following RCT

- Superiority trial
- Comparing new vs standard treatment
- Uses a composite primary endpoint consisting of CV death and hospitalisation for chronic heart failure (CHF)
- 1:1 randomisation

- Recognises the (obvious) fact that CV death is a more important event than CHF hospitalisation
- Recognises that patients have widely different risk profiles
Details of the New Method (2)

- Comparing patients on the new treatment vs patient on standard treatment
  - Did either one have a CV death before the other?
  - If neither had a CV death which patient had a CHF hospitalisation first?
- Essence of the approach
- Risk stratification is used to account for the underlying risk of the composite endpoint
  - Avoids unfair comparisons
  - Facilitates the creation of matched pairs
Details of the New Method (3)

• Matched Pairs Approach
  1. Matched pairs are formed
    • Matching method will vary – recommend individual patient risk matching
  2. For each pair determine which of the following categories the pair falls into:
    a. New patient had CV death first
    b. New patient has CHF hosp first
    c. Neither of a, b, d or e
    d. Standard patient had CHF hosp first
    e. Standard patient had CHF hosp first
  3. Calculate $N_a$, $N_b$, $N_c$, $N_d$ & $N_e$ the number of matched pairs in each category
How Outcomes Are Determined

- **A wins on death**
  - A
  - B

- **A wins on hospitalisation**
  - A
  - B

- **Tied or no winner**
  - A
  - B

Legend:
- **Death**
- **Hospitalisation**
- **Censored**
Details of the New Method (5)

- Trial findings can be summarised using the five values of $N_a$, $N_b$, $N_c$, $N_d$ & $N_e$
- The treatment difference can be summarised by the win ratio estimator
- The win ratio for the new treatment is calculated as follows
  - $N_w = N_d + N_e$ are the number of “winners” for the new treatment
  - $N_L = N_a + N_b$ are the number of “losers” for the new treatment
  - The summary estimate is $R_w = N_w / N_L$ is the “win ratio”
The CI and P-value are then calculated as follows

- A CI for the proportion winning $p_W = \frac{N_w}{(N_W + N_L)}$ is calculated using the regular CI for a binary proportion
  - $p_W \pm 1.96\left(\frac{p_W(1-p_W)}{(N_W + N_L)}\right)^{1/2} = (p_L, p_U)$
- This is then transformed back to a CI for $R_w$
  - $p_L/ (1-p_L), p_U/ (1-p_U)$

A test of the hypothesis of $R_w = 1$ is equivalent to a test of the binomial proportion $p_W = \frac{1}{2}$

- $Z = \frac{(p_W - \frac{1}{2})}{\left(\frac{p_W(1-p_W)}{(N_W + N_L)}\right)^{1/2}}$ where $Z$ is a standardised normal deviate under the null hypothesis
Details of the New Method (7)

• Unmatched Approach
  • In some cases matching may not be possible so an unmatched version of the approach is as follows
  • Let \( N_n \) and \( N_s \) be the number of patients in the new and standard treatment groups respectively
    1. Perform all possible \( N_n \times N_s \) pair wise comparisons
    2. As for the matched analysis each pair is placed into one of categories a, b, c, d, or e
    3. Calculate \( N_a, N_b, N_c, N_d \) & \( N_e \) the number of all pairs in each category except here
       • \( N_a + N_b + N_c + N_d + N_e = N_n \times N_s \)
       • As before the summary estimate \( R_W \) is computed
Details of the New Method (8)

• Unmatched Approach
  • No simple method exists to find the CI for $R_w$
    • We computed a CI using the Bootstrap
  • Finklestein and Schoenfeld (Stats in Medicine, 1999) developed a non-parametric significance test for the unmatched case
Example: EMPHASIS HF

- Compared eplerenone v placebo in 2737 patients with NYHA class II HF and ejection fraction ≤ 35%
- Primary outcome was a composite of CV health or HF hosp
- Reported results:
  - 21 months median follow up
  - 18.3% v 25.9% primary outcome in the eplerenone and placebo groups respectively
  - Hazard ratio 0.63 (95% CI 0.54 – 0.74), P = 0.002
- Issue with conventional analysis is the HF hosp tend to occur first and so the impact of eplerenone on CV death is lost in the composite
Example: EMPHASIS HF

• The new approach was applied in three ways

1. Matched Pairs
   • A risk score was built based on the 9 baseline variables in the original Cox model (excluding treatment)
   • Due to unequal sample size in the eplerenone and placebo groups 9 patients were dropped from the placebo group
   • Patients were then ranked from highest risk to lowest risk in each group based on the risk score
   • Each patient in the eplerenone group was then paired off against the same ranked patient in the placebo group
## Results: EMPHASIS HF

<table>
<thead>
<tr>
<th>Event Description</th>
<th>Matched Pairs</th>
<th>Matched Pairs Time Stratified</th>
<th>All Unmatched Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) CV death on eplerenone first</td>
<td>90</td>
<td>105</td>
<td>124825</td>
</tr>
<tr>
<td>b) HF hosp on eplerenone first</td>
<td>61</td>
<td>60</td>
<td>86127</td>
</tr>
<tr>
<td>c) none of the other 4 categories</td>
<td>964</td>
<td>914</td>
<td>1323085</td>
</tr>
<tr>
<td>d) HF hosp on placebo first</td>
<td>131</td>
<td>137</td>
<td>175606</td>
</tr>
<tr>
<td>e) CV death on placebo first</td>
<td>118</td>
<td>148</td>
<td>163129</td>
</tr>
<tr>
<td>Total No. of Pairs</td>
<td>1364</td>
<td>1364</td>
<td>1872772</td>
</tr>
</tbody>
</table>

Win ratio for composite event with 95% CI and z-score:

- Win ratio: 1.65 (1.35, 2.03) z-score: 5.05
- Win ratio: 1.73 (1.43, 2.10) z-score: 5.87
- Win ratio: 1.61 (1.37, 1.89) z-score: 5.45

Win ratio for CV death only with 95% CI and z-score:

- Win ratio: 1.31 (1.00, 1.74) z-score: 1.96
- Win ratio: 1.41 (1.10, 1.82) z-score: 2.74
- Win ratio: 1.31 (1.04, 1.66) z-score: 2.25
How Outcomes Are Determined

- **A wins on death**
  - A
  - B
  - A
  - B
  - A
  - B

- **A wins on hospitalisation**
  - A
  - B
  - A
  - B
  - A
  - B

- **Tied or no winner**
  - A
  - B
  - A
  - B

Legend:
- Death
- Hospitalisation
- Censored
2. Matched Pairs, Time Stratified
   • Patients were grouped into five equal strata based on their randomisation date
   • Patients were then risk matched within each time strata
   • This was an attempt to avoid “wasting” events where patients with very different follow up times were paired
   • Using time stratification reduced the number of “wasted” CV deaths from 68 to 30, and “wasted” CHF hospitalisations from 82 to 38 compared to the matched pair analysis

Example: EMPHASIS HF
### Results: EMPHASIS HF

<table>
<thead>
<tr>
<th>Category</th>
<th>Matched Pairs</th>
<th>Matched Pairs Time Stratified</th>
<th>All Unmatched Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) CV death on eplerenone first</td>
<td>90</td>
<td>105</td>
<td>124825</td>
</tr>
<tr>
<td>b) HF hosp on eplerenone first</td>
<td>61</td>
<td>60</td>
<td>86127</td>
</tr>
<tr>
<td>c) none of the other 4 categories</td>
<td>964</td>
<td>914</td>
<td>1323085</td>
</tr>
<tr>
<td>d) HF hosp on placebo first</td>
<td>131</td>
<td>137</td>
<td>175606</td>
</tr>
<tr>
<td>e) CV death on placebo first</td>
<td>118</td>
<td>148</td>
<td>163129</td>
</tr>
<tr>
<td><strong>Total No. of Pairs</strong></td>
<td>1364</td>
<td>1364</td>
<td>1872772</td>
</tr>
</tbody>
</table>

**Win ratio for composite**

<table>
<thead>
<tr>
<th></th>
<th>Matched Pairs</th>
<th>Matched Pairs Time Stratified</th>
<th>All Unmatched Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>(95% CI)</td>
<td>(1.35, 2.03)</td>
<td>(1.43, 2.10)</td>
<td>(1.37, 1.89)</td>
</tr>
<tr>
<td>z-score</td>
<td>5.05</td>
<td>5.87</td>
<td>5.45</td>
</tr>
</tbody>
</table>

**Win ratio for CV death only**

<table>
<thead>
<tr>
<th></th>
<th>Matched Pairs</th>
<th>Matched Pairs Time Stratified</th>
<th>All Unmatched Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>(95% CI)</td>
<td>(1.00, 1.74)</td>
<td>(1.10, 1.82)</td>
<td>(1.04, 1.66)</td>
</tr>
<tr>
<td>z-score</td>
<td>1.96</td>
<td>2.74</td>
<td>2.25</td>
</tr>
</tbody>
</table>
3. Unmatched Pairs
   
   • Each patient in the eplerenone group was compared with every patient in the placebo group
   
   • This leads to $1364 \times 1373 = 1872772$ unmatched pairs
Results: EMMPHASIS HF

<table>
<thead>
<tr>
<th>Category</th>
<th>Matched Pairs</th>
<th>Matched Pairs Time Stratified</th>
<th>All Unmatched Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) CV death on eplerenone first</td>
<td>90</td>
<td>105</td>
<td>124825</td>
</tr>
<tr>
<td>b) HF hosp on eplerenone first</td>
<td>61</td>
<td>60</td>
<td>86127</td>
</tr>
<tr>
<td>c) none of the other 4 categories</td>
<td>964</td>
<td>914</td>
<td>1323085</td>
</tr>
<tr>
<td>d) HF hosp on placebo first</td>
<td>131</td>
<td>137</td>
<td>175606</td>
</tr>
<tr>
<td>e) CV death on placebo first</td>
<td>118</td>
<td>148</td>
<td>163129</td>
</tr>
</tbody>
</table>

Total No. of Pairs: 1364

Win ratio for composite:
- 1.65 (95% CI: 1.35, 2.03)
- 1.73 (95% CI: 1.43, 2.10)
- 1.61 (95% CI: 1.37, 1.89)
- z-score: 5.05
- z-score: 5.87
- z-score: 5.45

Win ratio for CV death only:
- 1.31 (95% CI: 1.00, 1.74)
- 1.41 (95% CI: 1.10, 1.82)
- 1.31 (95% CI: 1.04, 1.66)
- z-score: 1.96
- z-score: 2.74
- z-score: 2.25
Summary

• There is a need to re-think the way composite endpoints are utilised in major trials
• The win ratio provides an enterprising and clinically relevant of giving priority to the more major components of any composite, e.g., mortality
• The win ratio is conceptually simple and straightforward to apply
Reference

Questions?