An experience in implementing the Promising Zone sample size re-estimation methodology in a phase 3 oncology study

Nikhil Chauhan, Director Biostatistics, BTG
Forward-looking statements

This presentation and information communicated verbally to you may contain certain projections and other forward-looking statements with respect to the financial condition, results of operations, businesses and prospects of BTG plc (“BTG”). These statements are based on current expectations and involve risk and uncertainty because they relate to events and depend upon circumstances that may or may not occur in the future. There are a number of factors which could cause actual results or developments to differ materially from those expressed or implied by these forward-looking statements. Any of the assumptions underlying these forward-looking statements could prove inaccurate or incorrect and therefore any results contemplated in the forward-looking statements may not actually be achieved. Nothing contained in this presentation or communicated verbally should be construed as a profit forecast or profit estimate. Investors or other recipients are cautioned not to place undue reliance on any forward-looking statements contained herein. BTG undertakes no obligation to update or revise (publicly or otherwise) any forward-looking statement, whether as a result of new information, future events or other circumstances. Neither this presentation nor any verbal communication shall constitute an invitation or inducement to any person to subscribe for or otherwise acquire securities in BTG.
Outline

- Study background
- Design options
- Selection of sample size re-estimation method
- Promising zone method
- Sample size calculations
- Regulatory authority interactions
Study background

- Device study, in oncology
  - Randomised control trial
  - Open-label
  - Time-to-event primary endpoint
  - Primary objective to obtain FDA marketing approval

- Study was ongoing
  - In early stages
  - No adaptive features

- Uncertainty in estimated treatment effect size
  - Published literature suggests treatment effect size may depend on aetiology, and other aspects of the disease
Study design options

• **Option 1**: Leave as is
  – But what if estimated treatment effect size is incorrect?
  – Considered too risky

• **Option 2**: Reduce treatment effect size and include interim analyses
  – Up-front commitment to large study
  – Considered to provide insufficient flexibility

• **Option 3**: Maintain treatment effect size and include unblinded sample size re-estimation
  – Predicted enrolment rates examined to ensure enrolment would not need to be suspended
  – Considered to offer sufficient flexibility
Sample size re-estimation methods

• **Solutions to preserve Type I error**
  – Combination of p-values before and after interim analysis
    • Pre-specify combination function
  – Preserve conditional Type I error
    • Set conditional Type I error for re-designed trial = conditional Type I error of current study design
  – Weighted statistics (Cui, Hung and Wang, 1999)
    • Down weights contribution of sample after interim analysis
      – Contradicts premise that “all patients are equal”

• **Overview**
  
  – Size of increase in sample size based on interim analysis results
  
  – Allowed only when interim results fall into the ‘Promising zone’
    
    • Promising zone is defined based on conditional power
  
• **Advantages**
  
  – Uses conventional test statistics
  
  – Equally weighted observations
  
  – Sample size increased only when interim results are ‘promising’
    (cf Option 2)
  
• **Disadvantage**
  
  – Spuriously positive interim analysis results could lead to sample size not being increased, resulting in an underpowered study

• **Type I error preservation**

  – Let $CP_1 = P_{Z_1}(Z_2 > c_2 \mid Z_1 = z_1)$
    
    • be the conditional power at first interim analysis
    
    • conditional probability of rejecting the null hypothesis at the final analysis given the interim results and the pre-planned sample size

  – For $CP_1 \geq 0.5$, the sample size may be increased and final analysis conducted using the conventional statistic, without inflating Type I error

  – In fact, Type I error preservation holds for $CP_1$ slightly less than 0.5, (e.g., $\geq$ a lower boundary = $CP_{\text{min}}$), in a design specific manner
### Design Adaptation Zones

<table>
<thead>
<tr>
<th>Condition</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfavourable: ( CP_1 &lt; CP_{min} )</td>
<td>Maintain sample size</td>
</tr>
<tr>
<td><strong>Promising:</strong> ( CP_{min} &lt; CP_1 &lt; 1 - \beta )</td>
<td>Increase sample size</td>
</tr>
<tr>
<td>Favourable: ( CP_1 \geq 1 - \beta )</td>
<td>Maintain sample size</td>
</tr>
</tbody>
</table>
Promising Zone method

Figure 1. Partitioning of Interim Result into Zones (†) and % Sample Size Increase in Each Zone: An Illustrative Example where \( n_{\text{max}}/n_2 = 2 \), \( n_1/n_2 = 0.5 \), one-sided \( \alpha = 0.025 \), and \( 1 - \beta = 0.9 \).

Some slight modifications
- Decision to only allow ‘fixed’ sample size increase
  - Rather than ‘sliding scale’ approach
  - Due to possibility of treatment effect size being back-calculated
- Inclusion of non-binding futility boundary
- Sample size re-estimation performed at second interim analysis

Design Adaptation Zones at 2nd interim analysis

<table>
<thead>
<tr>
<th>Zone</th>
<th>Conditions</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Futility</td>
<td>$CP_2 &lt; CP_{fut}$</td>
<td>Stop for futility</td>
</tr>
<tr>
<td>Unfavourable</td>
<td>$CP_{fut} &lt; CP_2 &lt; CP_{min}$</td>
<td>Maintain sample size</td>
</tr>
<tr>
<td>Promising</td>
<td>$CP_{min} &lt; CP_2 &lt; 1 - \beta$</td>
<td>Increase sample size</td>
</tr>
<tr>
<td>Favourable</td>
<td>$CP_2 \geq 1 - \beta$</td>
<td>Maintain sample size</td>
</tr>
<tr>
<td>Efficacy</td>
<td>Efficacy boundary crossed</td>
<td>Stop for efficacy</td>
</tr>
</tbody>
</table>
Sample size calculations

Computations performed in SAS PROC SEQDESIGN and EAST

Timing of interim analyses and alpha spending options discussed with team

– Based on predicted enrolment rates to ensure enrolment would not need to be suspended

– Sufficient power to stop early for efficacy if treatment effect is large

Power family alpha spending methods considered

– O’Brien-Fleming ($\rho = 0.5$)

– Pocock ($\rho = 0$)

– Other values of $\rho$

<table>
<thead>
<tr>
<th></th>
<th>OBF</th>
<th>$\rho=0.25$</th>
<th>Pocock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of events</td>
<td>202</td>
<td>208</td>
<td>222</td>
</tr>
<tr>
<td>Number of patients</td>
<td>332</td>
<td>342</td>
<td>365</td>
</tr>
<tr>
<td>Timing</td>
<td>46</td>
<td>47</td>
<td>47</td>
</tr>
<tr>
<td>Power for $\Delta=6\text{mo}$</td>
<td>0.55</td>
<td>0.61</td>
<td>0.66</td>
</tr>
<tr>
<td>Power for $\Delta=8\text{mo}$</td>
<td>0.68</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>Power for $\Delta=10\text{mo}$</td>
<td>0.79</td>
<td>0.83</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Numbers in the table above are for illustration purposes only
FDA interactions

• Protocol amendment submitted to FDA
  – Conditional approval
    • To mitigate the introduction of operational bias into this open-label study, please have operational procedures in place to ensure that the IDMC does not reveal unblinded interim results to study investigators per the recommendations given in Mehta and Pocock (Stat. Med., 2011).

• You need to specify your adaptive method more clearly and justify it. For example, with your method, what values for the conditional power cause the sample size to increase or not increase? Also do the similarities between your method and Mehta and Pocock’s method imply that your method controls the type 1 error?
• IDMC charter includes the following to minimize operational bias:
  – Unblinded CRO statistician and statistical programmer(s), who are not otherwise be involved in the study, will perform the interim analyses
  – IDMC members required to maintain strict confidentiality of study data
  • Not share any study data or information about the study with any individual external to the IDMC, including study investigators and Sponsor staff involved in operational aspects of the study
  – Fixed sample size increase allowed (vs ‘sliding scale’ approach)
  – Procedure mapped out for notifications to be followed by IDMC if recommendation to stop study or to increase sample size
FDA interactions

- Responses to FDA

Enrol Subjects

Stop for Futility
CP₁ ≤ %

First Interim Analysis
at Events

Stop for Efficacy if efficacy boundary crossed

Stop for Futility
CP₂ ≤ %

Second Interim Analysis
at events

Stop for Efficacy if efficacy boundary crossed

Unfavorable Zone:
CP₂ < %
events for final analysis

Promising Zone
% ≤ CP₂ < %
... events for final analysis

Favorable Zone
CP₂ ≥ %
events for final analysis

CP₁ = Conditional power at the first interim analysis; CP₂ = Conditional power at the second interim analysis
FDA interactions

- Theoretical explanation of control of Type I error and derivation of $\text{CP}_{\text{min}}$

![Diagram showing zones and conditional power](image)
FDA interactions

- Simulation study also conducted
  - Type I error assessment
    - Assuming hazard ratio = 1 and nominal one-sided alpha = 2.5%

<table>
<thead>
<tr>
<th>Zone</th>
<th>Prob. of entering each zone (%)</th>
<th>Prob. of declaring efficacy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group Sequential Design</td>
<td>Promising Zone Design</td>
</tr>
<tr>
<td>Futility</td>
<td>89.9</td>
<td>0</td>
</tr>
<tr>
<td>Unfavourable</td>
<td>3.8</td>
<td>4.5</td>
</tr>
<tr>
<td>Promising</td>
<td>4</td>
<td>11.9</td>
</tr>
<tr>
<td>Favourable</td>
<td>1.1</td>
<td>27.2</td>
</tr>
<tr>
<td>Efficacy</td>
<td>1.2</td>
<td>100</td>
</tr>
<tr>
<td>All Trials</td>
<td>100</td>
<td>2.133</td>
</tr>
</tbody>
</table>

- Average sample sizes, number of events, study durations also provided
Simulation study

- Type II error assessment also provided to FDA

- Assuming hazard ratios corresponding to both original and increased sample sizes

<table>
<thead>
<tr>
<th>Zone</th>
<th>Prob. of entering each zone (%)</th>
<th>Prob. of declaring efficacy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Group Sequential Design</td>
</tr>
<tr>
<td>Futility</td>
<td>13.6</td>
<td>0</td>
</tr>
<tr>
<td>Unfavourable</td>
<td>4.7</td>
<td>54.8</td>
</tr>
<tr>
<td>Promising</td>
<td>15.3</td>
<td>74.5</td>
</tr>
<tr>
<td>Favourable</td>
<td>12.5</td>
<td>88.6</td>
</tr>
<tr>
<td>Efficacy</td>
<td>53.9</td>
<td>100</td>
</tr>
<tr>
<td>All Trials</td>
<td>100</td>
<td>82.2</td>
</tr>
</tbody>
</table>

Numbers in the table above are for illustration purposes only.
Conclusions

- Sample size re-estimation offered by the promising zone method fitted needs of BTG
- Positive experience with device branch of FDA
  - FDA agreement to change design of an ongoing, open-label study
- Able to show control of Type I error theoretically, but simulation study gave additional assurance
- Our study investigators tell us they like the “Promising Zone” concept!
Questions