End Points

“Although the final statistical analysis plan indicated that the mean time to the first fracture would be reported as the primary outcome, because the overall rate of fractures was low, the hazard ratio for fracture was... reported as the primary outcome.”

- Mean time to first fracture
  - Why did they originally choose???
  - Large percentage (> 80%) have no fractures within 2 years
  - (Right) Censored: Subjects not observed long enough for fracture to occur

- Survival analysis: Good analysis option
  - Kaplan-Meier, Cox proportional hazards model
  - Models time-to-event data with censoring
  - Tests if two curves (e.g. Fig 2) are the same
Multiple Comparisons

- **Goal:** Test multiple hypotheses while controlling $\alpha$
  - $\alpha$: Significance level (Type I error rate)
- **General Method:** Bonferroni correction
  - Divide $\alpha$ by the number of tests performed (e.g. $\alpha/3$)
  - Conservative: Assumes all tests are independent
- **Interim Analyses**
  - Test one hypothesis multiple times, but preserve $\alpha$
  - Stop trial early if...
    - Zoledronic acid very effective
    - Many adverse events
- **O’Brien-Fleming**
  - Method for multiple comparisons with interim analyses
  - “Spend” most of $\alpha$ at end of trial
Test statistics ($Z$) and p-values ($p$)

- $Z$: Test statistic (follows a Normal distribution)
- Rejecting the null hypothesis...
  - ... if $p$ is small (e.g. $p < .05$)
  - ... if $Z$ is large (e.g. $Z > 1.96$)

Figure: Sequential outcomes and boundaries for interim standardized test statistics from a clinical trial.
### O'Brien-Fleming Bounds

#### Three Interim Analyses

<table>
<thead>
<tr>
<th>Time</th>
<th>Spent α</th>
<th>Cum. α</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>0.75</td>
<td>0.016</td>
<td>0.019</td>
</tr>
<tr>
<td>1.00</td>
<td>0.031</td>
<td>0.050</td>
</tr>
</tbody>
</table>

The graph shows the O'Brien-Fleming bounds over time for three interim analyses. The critical Z statistic values are indicated by the lines and markers on the graph, with time on the x-axis and Z statistic on the y-axis. The α levels and cumulative α levels are shown in the table above. The graph illustrates the decision points for rejecting or failing to reject the null hypothesis at each analysis point.
O’Brien-Fleming and Bonferroni

Three Interim Analyses

<table>
<thead>
<tr>
<th>Time</th>
<th>O’Brien-Fleming</th>
<th>Bonferroni</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>0.75</td>
<td>0.016</td>
<td>0.019</td>
</tr>
<tr>
<td>1.00</td>
<td>0.031</td>
<td>0.050</td>
</tr>
</tbody>
</table>

- O’Brien-Fleming: Reject Null
- Bonferroni: Fail to Reject Null