Endpoint Selection in Phase II Oncology trials

Anastasia Ivanova

Department of Biostatistics
UNC at Chapel Hill
aivanova@bios.unc.edu
Primary endpoints in Phase II trials

Recently looked at journal articles published in leading oncology journals in 2005 and 2010

- Country
- Cancer site
- Single center vs multicenter
- Design (single stage, Simon, Fleming etc)
- Number of arms (1 or 2),
  if more than 1 arm, was it randomized?
- Primary endpoint, secondary endpoints
- Was \( H_0 \) rejected? Was new treatment recommended?
- Were stats clearly described?
Primary endpoints in Phase II trials

• Statistical design

<table>
<thead>
<tr>
<th>Design</th>
<th>2005</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single stage</td>
<td>32% of articles</td>
<td>46% of articles</td>
</tr>
<tr>
<td>Simon’s</td>
<td>58% of articles</td>
<td>46% of articles</td>
</tr>
<tr>
<td>Fleming’s</td>
<td>8% of articles</td>
<td>8% of articles</td>
</tr>
</tbody>
</table>
Primary endpoints in Phase II trials

- Size of the trial

<table>
<thead>
<tr>
<th>Size of Trial</th>
<th>2005</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 40</td>
<td>74% of articles</td>
<td>48% of articles</td>
</tr>
<tr>
<td>41-60</td>
<td>18% of articles</td>
<td>26% of articles</td>
</tr>
<tr>
<td>&lt; 60</td>
<td>8% of articles</td>
<td>26% of articles</td>
</tr>
</tbody>
</table>
Primary endpoints in Phase II trials

- Statistical design is **NOT** clearly described

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>18%</td>
</tr>
<tr>
<td>2010</td>
<td>5%</td>
</tr>
</tbody>
</table>
Primary endpoints in Phase II trials

- Response rate (RR) is the most frequent primary endpoint in Phase II oncology trials (preliminary results)

Response \(\text{(RR)}\) = Complete Response + Partial Response

<table>
<thead>
<tr>
<th>Year</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>81% of articles</td>
</tr>
<tr>
<td>2010</td>
<td>54% of articles</td>
</tr>
</tbody>
</table>
PFS and PFS6

• The second most frequently used primary endpoint is

  PFS (Progression Free Survival) 15%
  PFS6, PFS3, PFS2, PFS1.5 15%

• Trials with PFS6 use Simon’s design to allow for early stopping for futility.

• Data are summarized using a Kaplan-Meier curve rather than proportion of patients with PFS6
CR, PR and SD

• Other primary endpoints include
  - **Complete Response** (CR)
  - **Disease control** = CR + PR + Stable Disease
  - **Clinical benefit** = CR + PR + Stable Disease >3 months
Growth modulation index (GMI)

- Each patient receives **two successive treatments**, standard first and then experimental (usually a biomarker driven)

- $\text{TTP}(1)$ - time to progression on the first treatment
  $\text{TTP}(2)$ - time to progression on the second treatment

- $\text{GMI} = \frac{\text{TTP}(2)}{\text{TTP}(1)}$
Growth modulation index (GMI)

• GMI was introduced by Von Hoff (1998)
• Von Hoff et al. (2010)

• Cousin (2013) showed that GMI of higher than 1.33 was associated with significant improvement in OS
## Proportion of successful Phase II trials

<table>
<thead>
<tr>
<th></th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null hypothesis rejected</td>
<td>82%</td>
<td>18%</td>
</tr>
<tr>
<td>Drug recommend for future</td>
<td>44%</td>
<td>56%</td>
</tr>
<tr>
<td>investigation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Successful Phase II trials

- If $H_0$ was not rejected based on RR, however PFS or OS was good, and therefore the drug is recommended

- Publications with stats not clearly described were significantly more likely to recommend the drug for further study
Tumor Response versus Disease Control

CR = Complete response
PR = Partial Response
SD = Stable Disease
PD = Progressive Disease
Tumor Response (TR) = CR + PR

Disease Control (DC) = TR + SD
Tumor Response versus Disease Control

• Historically the RR was a frequently used endpoint to assess novel cytotoxic agents which induced tumor regression

• Many targeted agents are *cytostatic* and inhibit tumor growth rather than cause tumor regression.

• In situations such as this, when RR is modest, clinicians may be interested in disease control (DC) since it may better predict clinical outcome.
Tumor Response versus Disease Control

• An unblinded single arm trial may be susceptible to bias in the assessment of tumor response by the investigator and lacks a “control arm” for comparison.

• Issues in determining tumor response have included variation in the assessment of response or variability in tumor measurements on repeat imaging.
Example: Phase II trial in NSCLC, testing TR and DC simultaneously

Phase II trial in patients with advanced non-small cell lung cancer who have experienced disease progression after platinum-based therapy
Principal Investigator Tom Stinchcombe

Simon’s minimax design for DC
\( p_{0D} = 0.25, \ p_{1D} = 0.50, \ \alpha = 0.05 \)
24 patients are needed for power of 80%

Simon’s minimax design for RR
\( p_{0T} = 0.05, \ p_{1T} = 0.25, \ \alpha = 0.05 \)
16 patients are needed for power of 80%
Testing TR and DC (TR + SD) simultaneously

Consider two hypotheses:
about tumor response (TR) rate
and about disease control (DC) rate, DC = TR + SD

\[ H_{0T} : \ p_T < p_{0T} \quad \text{versus} \quad H_{1T} : \ p_T > p_{0T} \]
\[ H_{0D} : \ p_D < p_{0D} \quad \text{versus} \quad H_{1D} : \ p_D > p_{0D} \]

We would like to test \( H_{0T} \cap H_{0D} \)

The drug is accepted if TR rate is good OR DC rate is good, that is, if one of the individual hypothesis is rejected.
The drug is rejected if BOTH TR rate and DC rate are poor.
Example. \( p_{0T} = 0.15, \ p_{1T} = 0.55, \ p_{0D} = 0.35, \ p_{1D} = 0.75, \ \alpha = 0.05 \)
Example: Phase II trial in NSCLC, testing TR and DC simultaneously

Design used in the trial

\[ p_{0T} = 0.05, \quad p_{1T} = 0.25, \]
\[ p_{0D} = 0.25, \quad p_{1D} = 0.50, \quad \alpha = 0.05 \]
\[ n = 18, \quad n_{1} = 12 \]

Power for \( H_{0T} \cap H_{0D} \) under various alternatives

<table>
<thead>
<tr>
<th>BOTH</th>
<th>TR only</th>
<th>DC only</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{1T} ) = 0.25</td>
<td>( p_{1T} ) = 0.25</td>
<td>( p_{1T} ) = 0</td>
</tr>
<tr>
<td>( p_{1D} ) = 0.50</td>
<td>( p_{1D} ) = 0.25</td>
<td>( p_{1D} ) = 0.50</td>
</tr>
<tr>
<td>0.80</td>
<td>0.76</td>
<td>0.70</td>
</tr>
</tbody>
</table>
Example: A Phase II trial with relaxed stopping for futility

A Phase II trial of a chemotherapy and a monoclonal antibody in older patients with previously untreated diffuse large B-cell lymphoma

Principal Investigator  Steven I. Park

- **Tumor Response (TR) is the primary outcome**
  Would like to use Simon’s two-stage design with TR as the primary endpoint.

- **Would like to base futility stopping on the rate of Disease Control**
  \[ DC = TR + \text{Stable Disease} \], at the end of the trial
Interim analysis for futility
## Phase II oncology trials

<table>
<thead>
<tr>
<th>Simon (1989)</th>
<th>Extensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single arm</td>
<td>Two-arm (Kepner, 2011)</td>
</tr>
<tr>
<td>Two-stage</td>
<td>Three-stage (Cheng, 2007)</td>
</tr>
<tr>
<td>Early stopping for futility only</td>
<td>Early stopping for futility and efficacy (Fleming, 1982)</td>
</tr>
<tr>
<td>min((N)) or (\min(E_0 N))</td>
<td>Admissible (Jung et al., 2004)</td>
</tr>
<tr>
<td>Binary endpoint</td>
<td>Ordinal outcome</td>
</tr>
<tr>
<td></td>
<td>Lu et al. (2005), Ivanova et al. (2012)</td>
</tr>
</tbody>
</table>
Software for Phase II oncology trials

Easy to use web-based software
http://cancer.unc.edu/biostatistics/program/ivanova/

- Simon’s admissible designs
- Fleming’s admissible designs
- Two-stage designs with relaxed stopping for futility
- Two-stage designs with ordinal outcome
- Continuous toxicity monitoring

- Rapid Enrollment Design (RED) for dose finding (coming soon)
References


Fleming, TR (1982). One-Sample multiple testing procedure for Phase II clinical trials, *Biometrics* 38: 143-151


