Phase II Clinical Trials

Richard L. Schilsky, M.D.
Professor of Medicine
Chief, Section of Hematology/Oncology
University of Chicago
An Approach to Classifying Clinical Studies According to Objective

- Phase I: Human Pharmacology
- Phase II: Therapeutic Exploratory*
- Phase III: Therapeutic Confirmatory
- Phase IV: Therapeutic Use

Federal Register 62, No. 242 1997
Goals in Phase II

- Estimate activity
- Describe toxicity
- Develop PK-PD correlations
- Confirm drug-target interactions
Essential Elements of Phase II Trials

- Why?
- Who?
- What?
- How?
Essential Elements of Phase II Trials

Why?

Objective and hypothesis

• State clearly one primary hypothesis-testing objective

• May include a limited number of secondary objectives
Primary Objective of a Phase II Trial

• Provide an estimate of the clinical “activity” of a new treatment approach:

• Examples:
  – To determine the objective response rate (CR + PR) of drug A in patients with advanced X cancer
  – To determine the 6 month progression-free survival (PFS) rate of the combination AB in patients with recurrent or metastatic Y cancer
Phase II Design Issues

- Endpoint: response rate vs other
- Level of interest
- Precision of estimate
- Making the go/ no go decision to move to phase III
Classical Phase II Design

- Endpoint: objective response (need to prospectively define response criteria)
- Two stage design typical unless accrual expected to be very rapid and/or toxicity low
- Primary ethical dilemma: continued accrual in the absence of responses
Stage I

- True response rate must be less than X% if no responses seen in first n patients

<table>
<thead>
<tr>
<th>True response rate</th>
<th>5%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>49</td>
<td>29</td>
<td>14</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>10%</td>
<td>45</td>
<td>22</td>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>
Stage II

• Final sample size depends on desired level of activity and error limits
• Alpha error: probability of accepting an inactive drug
• Beta error: probability of rejecting an active drug
• Objective RR may not be most important endpoint
• Consider CR rate; pCR; response duration
Secondary Objective(s) of a Phase II Trial

• Examples:
  – Hypothesis-generating
  – Subset analysis
  – Alternative endpoint analysis (e.g. survival)
  – Correlative science
  – Better define toxicity in a more homogeneous population
Essential Elements of Phase II Trials

Who?

Patient selection

• Inclusion and exclusion criteria
Choice of Patient Population - 1

• Disease:
  – Be explicit
  – e.g. malignant neuroendocrine tumors: Do you include all subtypes? well-differentiated carcinoid tumors have different biology compared to small cell variants

• Target expression:
  – Is there a biological basis to enrich the patient population?
  – Is the turn-around time for the screening procedure fast enough to be feasible?
Choice of Patient Population - 2

• Performance status
• Organ function
• Prior therapy allowed:
  – What prior systemic therapy is allowed?
  – For targeted therapies, should exclude other targeted therapies of the same class if cross-resistance is a concern
  – Should exclude prior extensive radiotherapy if myelosuppression is a concern
• Concomitant therapy disallowed:
  – Concern for drug interactions e.g. Cyt P450
Essential Elements of Phase II Trials

What?

- Drug, dose, schedule
- Specify supportive therapy allowed
- Dose modification
- Dosing details:
  - IV: premedications, infusion details
  - PO: fasting or fed, what to do if vomited, missed dose

Treatment plan
Essential Elements of Phase II Trials

Endpoints:
- Measurable tumor mass reduction
- Progression-based endpoints: TTP, PFS
- Serologic response: PSA, CA125
- Survival
- Disease “stabilization”
- Correlative studies

How? Endpoints and design
Measurable Tumor Mass Reduction

**Pro**
- “Standard”
- “Proves” antitumor activity
- Standard criteria
- Early readout

**Con**
- Not applicable in all cancers e.g. prostate
- Difficult to assess in some cancers e.g. mesothelioma
- “RR” relationship to survival (is RR a surrogate for an important effect?)
- May not be appropriate for some classes of compounds
- Response criteria arbitrary
Reported response rates in phase III and phase II studies


Copyright © American Society of Clinical Oncology
Assessing Activity if No Response Expected

- Absence of progression at specified time point
- Time to progression
- Both require randomized controls unless natural history very well defined
Progression-Based Endpoints

**Pro**
- Reflects “biologic activity”
- More clinically meaningful than RR for some agents

**Con**
- Variable definition (serologic, symptomatic, measurable disease)
- Influenced by frequency of observation
- Influenced by tumor doubling-time more than therapy
TTP Better Categorizes Tumor Control Than Response Rate

Response Status
- Progressive Disease
- Response
- Stable Disease

Time (weeks)
- 0
- 6
- 12
- 18
- 24
- 30
- 36
- 42
- 48
- 54

Total Target Tumor Length (cm)
- 0
- 10
- 20
- 30
- 40
- 50
- 60
- 70
- 80

PD at 6 wks
PD at 18 wks
PD at 54 wks
Time to Progression

• Precision depends on identification of all lesions at baseline and on frequency of evaluation
• Always an estimate since actual progression occurs between observations
• Requires control for rate of progression in absence of treatment effect
• Unblinded studies subject to ascertainment bias
Progression depends on when you look

@ 2 months: 50% progressed
Progression depends on when you look

@ 3 months: 100% progressed

Progressive Disease

Death

Change in Marker of Response

Day 1 of Study
Time to Progression/PFS

• Includes all patients in analysis
• Endpoint sooner than survival; no crossover effect
• Definition of progression
  - death due to cancer
  - new lesions
  - increase in size of existing lesions (?)
  - increase in tumor metabolism
  - increase in plasma level of tumor marker
  - decline in PS or increase in symptoms
Objective response to Ipilimumab after initial significant progression

Screening

Week 12: Progression

Week 20: Regression

Week 36: Still Regressing

Reproduced from Wolchok. ASCO. 2008 (abstr 3020)
Time to Progression

• How much improvement in TTP constitutes benefit to a patient, particularly one who is asymptomatic?
  - 2 months?
  - 4 months?
  - 6 months?
Survival

**Pro**
- Easy to measure!
- Easy to quantitate!

**Con**
- Delayed endpoint
- Influenced by “subsequent” therapies
Chemotherapy +/- Trastuzumab for Metastatic Breast Cancer: OS

2/3 of patients on chemo alone received trastuzumab post-progression

Stabilization

**Pro**

- Nothing!

**Con**

- Depends on intrinsic growth rate of the tumor …which is almost NEVER measured in a clinical trial

Unless you employ special design (randomized discontinuation) or you measure doubling-time prior to therapy
“Stabilization” also depends on the biology of the disease.

Progressive Disease

Day 1 of Study

Day 180 of Study
Design Options

- Single arm, 2 stage
- Randomized, phase II
- Randomized discontinuation
- Bayesian
Single-Arm, 2-Stage Design (Simon, Mini-max,..)

• Treat ~12-18 patients at 1st stage
• Determine the “response rate”
  • Less than that projected to indicate activity (p0): STOP!
  • Sufficiently great to indicate activity: CONTINUE
• At the end of 2nd stage, declare drug / intervention worthy of further evaluation if > x number of “responses” are observed (p1)
Example of Single-Arm, 2-Stage Design

- 2-stage (Simon) of gefitinib in NSCLC as first-line therapy
- $p_0 = 0.1$, $p_1 = 0.25$
- Stage I: proceed if $\geq 2$ PR in 22 pts
- Stage II: interest if $\geq 7$ PR in 40 pts
- Results:
  - 12 PR / 40 pts (RR = 30%, 95% C.I. 17% - 47%)

Niho et al, J Clin Oncol, 2006
Randomized Phase II Trial Designs

Sharma, et al, JNCI 2011; 103:1093

Randomized vs Control

Randomized Discontinuation

Delayed Start

Adaptive

Pick the Winner

Phase II-III
Types of Placebo-Controlled Randomized Phase II Trials

• Simple multiple dose study
  – All patients randomized between placebo and multiple doses (A, B, C, etc.)
  – After primary endpoint (e.g. progressive disease) has been met, patients go off study
  – If time to progression (TTP) greater in any drug vs. placebo group, drug is active

  • Would expect that drug would be active at all doses above some threshold
    – Toxicity may obscure effect at higher doses
Randomized phase II study of multiple dose levels of temsirolimus in patients with advanced refractory renal cell carcinoma

<table>
<thead>
<tr>
<th>CCI-779 mg</th>
<th>n</th>
<th>Median mos</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>36</td>
<td>6.3</td>
<td>3.6, 7.8</td>
</tr>
<tr>
<td>75</td>
<td>38</td>
<td>6.7</td>
<td>3.5, 8.5</td>
</tr>
<tr>
<td>250</td>
<td>37</td>
<td>5.2</td>
<td>3.7, 7.4</td>
</tr>
</tbody>
</table>

Log-Rank Test $P = .933$
Planned Crossover Designs

- Placebo vs. drug for fixed period of time followed by crossover
- Randomized discontinuation design with crossover
Placebo-controlled randomized trial followed by open-label treatment

- **Primary endpoint**: Progression-free rate over fixed period of time
- **Fixed duration randomized phase**
- **Open-label drug**
- **“Crossover” at fixed time point**
Randomized Phase II Design

1) Randomized phase II selection design:
   • “Pick-the-winner“
   • Randomized 2 or more experimental agents or regimens
   • “Winning” agent or regimen will then be subjected to a definitive phase III trial against standard
   • e.g. RP2T of carboplatin vs paclitaxel + carboplatin in platinum-sensitive recurrent advanced ovarian cancer (Gonzalez-Martin et al, Ann Oncol, 2005)
Randomized Phase II Design

2) Randomized phase II design including a reference standard treatment control arm:

- Non-comparative
- Standard control arm acts as a check
- Not to be compared directly with the experimental arm(s)
- e.g. RP2T of docetaxel + prednisone vs docetaxel + prednisone + new drug in hormone refractory prostate cancer
Randomized Phase II Design

3) Phase II/III trial design:

- Randomized phase II trial embedded within phase III trial
- This design is efficient only if the experimental regimen has a reasonably high likelihood of “success”, otherwise intensive efforts are expended in developing a phase III protocol
- Endpoint and operating characteristics of study are critical
Randomized Phase II Design

**Pro**
- Reduces selection bias
- Improve patient comparability
- Non-survival endpoints can be ascertained fairly quickly

**Con**
- Generally requires more pts than a single-arm phase II trial
- A “positive” result in the screening design often compromises the ability to conduct definitive phase III trials
Randomized Discontinuation Trials (RDT)

- Enrichment design, in which all pts receive study drug for an initial run-in period, followed by random assignment of patients with disease stabilization to either the study drug or placebo
RDT Design

Sorafenib 12-week run-in (n=202)

- Tumor shrinkage ≥25% (n=73)
- Tumor growth/shrinkage <25% (n=69)
- Tumor growth ≥25% (n=51†)

Continue open-label sorafenib (n=79)
- Placebo* 12 weeks (n=33)
- Sorafenib 12 weeks (n=32)
- Off study (n=58)

18% Progression free 24 weeks
50% Progression free 24 weeks

Disease status at 12 weeks unknown (n=9)

*Placebo patients who progressed could cross over to sorafenib
†Including 36 patients without bidimensional tumor measurements, but with radiological evidence of progression

Ratain et al, J Clin Oncol, 2006
12 Week Changes from Baseline in Bidimensional Radiographic Measurements*

Data shown for 166 patients (12-week bidimensional measurements were not available for 36 patients, who had radiographic evidence of progression). Mean change at 12 weeks = -18 ± 33%

*Investigator-assessed
Randomized Discontinuation Design

**Pro**
- ↓ heterogeneity of randomized pts
- ↑ statistical power
- Evaluates true disease “stabilization”
- Minimizes use of placebo

**Con**
- Large sample size
- Longer accrual time
- Many patients never reach stable disease
- Unknown how long to give a new agent
- No advantage unless high rate of disease stabilization
## Single-Arm versus Randomized Phase II Trials

<table>
<thead>
<tr>
<th>Design</th>
<th>Single-Arm Phase II</th>
<th>Randomized Phase II</th>
</tr>
</thead>
</table>
| Advantage| • Simple and easy to execute  
• Typically smaller sample size            | • Provides a contemporary control  
• Enhances collection of biospecimens to help development of predictive biomarkers |
| Disadvantage| • Historical control can be unreliable due to pt heterogeneity and changes in outcome based on supportive care and diagnostic techniques | • Larger sample size  
• Elevated $\alpha$ and $\beta$ error rates to reduce sample size can lead to false + or - results |

### Single-Arm versus Randomized Phase II Trials

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Single-Arm Phase II</th>
<th>Randomized Phase II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited pts available</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>Adequate historical data</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Single-agent therapy in pretreated pts with common tumor types</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Progression-based endpoints</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Addition of new agent to standard combination regimen</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Inadequate historical data</td>
<td>-</td>
<td>+++</td>
</tr>
</tbody>
</table>

Gan, Grothey, Pond, Moore, Siu, Sargent. *JCO*, 2010
Bayesian (Adaptive) Designs

- Uses prior information and assumptions in the decision making process
- Updates prior information with trial data
- Trial data will “overwhelm” the prior information in a well designed trial
- Outcome is a probability with “credible” intervals (e.g. the probability of RR >20% is 0.95)
Bayesian Design

**Pro**
- Both prior information and real trial data are incorporated in a thorough evaluation
- Can make frequent stop or go evaluations
- Collaboration with biostatisticians

**Con**
- Prior information is subjective
Biomarker based Trial Designs

• Traditional
  – All eligible patients
  – Assess overall treatment effect

• Enrichment design
  – Only patients with positive assays
  – Assess treatment effect in assay positive pts

• Hybrid design
  – All eligible patients
  – Assess overall and subgroup effect
Factors Affecting Efficiency of Biomarker Based Trials

- Prognostic vs. predictive marker
- Treatment effect in those with and those without the target
- Prevalence of target subgroup
- Accuracy of assay
- Cost and turnaround of assay
Effect of an Imperfect Assay

- False positives – decreases power by decreasing effect size
- False negatives – slower accrual by screening out patients who should have been entered
Conclusions

• Phase II trials are exploratory studies and rarely are definitive
• Efficient to exclude inactive therapies
• Results must be interpreted cautiously, in the context of the availability of other therapies
• Estimate clinical activity and provide further safety information – important in the “go/no go” decision
• Require confirmation in phase III trials