Bayesian Phase I/II clinical trials in Oncology

Pierre Mancini, Sandrine Micallef, Pierre Colin

Séminaire JEM-SFES - 26 Janvier 2012
Outline

- Oncology phase I trials
- Limitations of traditional phase I designs
- Bayesian phase I design with toxicity endpoint
- Bayesian phase I design with toxicity and efficacy
- I-SPY 2: example of adaptive phase II trial
- Bayesian adaptive phase III trials
- Conclusion
Phase I clinical trials in oncology

- Recommend a dose for Phase II clinical trial
- Design:
  - Patients included in successive cohorts (usually n=3 in each cohort)
  - All patients within the same cohort receive the same dose
    - First cohort receive the lowest dose
    - Primary endpoint: Dose-Limiting Toxicity
    - After completion of each cohort, decision is made on predefined algorithm to:
      - Escalate the dose
      - Stay at the same dose
      - De-escalate the dose
      - Stop the study
Original central hypothesis in cancer dose finding

- Therapeutic and toxic effect of a treatment are related to the dose given
- Monotonic dose-toxicity and dose-activity relationship
  - higher is the dose, higher is the activity
  - highly influenced oncologist in designing phase I trials
- True for cytotoxic drug but currently challenged for new generation of anti-cancer drug, e.g. targeted agents with less toxicity
Phase I purposes

- **Safe doses**
- **Active doses**

**Toxicity**
- Unacceptable Tox (e.g. proba of DLT)
- Unacceptable toxicity threshold: 33%

**Activity**
- Targeted effect. At least 60% of + response to PD marker

**Therapeutic interval**
- Minimum Effective Dose: defined on activity criterion
- Maximum Tolerated Dose: defined on a safety criterion

SANOFI
Algorithm-based ("3+3") phase I design

1. Treat 3 subjects
   - 0/3 DLT: Go to Higher dose
   - 1/3 DLT: Treat 3 more subjects
   - ≥2/3 DLT: Go to Lower Dose

   - 1/6 DLT: Go to Higher dose
   - 1/6 DLT: Treat 3 more
   - ≥2/6 DLT: Go to Lower Dose
   - 3 already treated: 6 already treated
   - 1/6 DLT: MTD
Simulation of 1000 phase I trials using “3+3” design

Distribution of estimated MTD

<table>
<thead>
<tr>
<th>Dose (DLT rate,%)</th>
<th>Doses</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;60 mg/m² (0%)</td>
<td>0.02</td>
</tr>
<tr>
<td>60 mg/m² (5%)</td>
<td>0.18</td>
</tr>
<tr>
<td>80 mg/m² (15%)</td>
<td>0.32</td>
</tr>
<tr>
<td>100 mg/m² (25%)</td>
<td>0.29</td>
</tr>
<tr>
<td>120 mg/m² (35%)</td>
<td></td>
</tr>
<tr>
<td>140 mg/m² (50%)</td>
<td>0.16</td>
</tr>
<tr>
<td>160 mg/m² (70%)</td>
<td>0.03</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Algorithm-based designs: Pros and Cons

● **Pros**
  ● Simplicity, Classical
  ● Generally « safe »

● **Cons**
  ● Short memory (only the current dose level used to decide about next one)
  ● High variability
  ● Tend to under-estimate MTD
  ● Too many pts treated at non-toxic (and non-active?) dose
    • but accelerated titration design better than « 3+3 »
  ● Choice of targeted toxicity level severely limited
What means « Dose-response model based » approach?

- Try to assess a dose-response relationship using mathematical function.
- Use mathematical tool (model) to define probability of DLT as a function of dose.

Mathematical model:

\[ \text{Tox} = f(\text{dose}, \text{parameter}) \]

- Provides quantification for the dose response relationship.
- Allows interpolation: « what happened between two dose levels? »
Knowledge after updating: Posterior Bayes Theory

Experimental data

Information anterior to the study

- Knowledge of clinicians
- Data from literature
- Information from other related studies
- Data from preclinical studies

A priori knowledge (expertise): Prior

Updated information on the basis of collected data

Knowledge after updating: Posterior

Current information of the study

Principle of the Bayesian approach
Estimated dose-response relationship: 
*a priori* and *a posteriori*

A priori

![Graph A priori]

A posteriori

![Graph A posteriori]
Phase I trial of Agent A + Agent B

- Chronology of escalating using the “3+3” design

```
Data
DLT: 0/3 11.5 mg/m²  →  DLT: 0/5 15.5 mg/m²  →  DLT: 0/3 20 mg/m²  →  DLT: 1/3 25 mg/m²  →  DLT: 0/3 25 mg/m²  →  DLT: 0/3 30 mg/m²  →  DLT: 0/3 35 mg/m²
```

What is the final estimated MTD?

“3+3” → 35 mg/m²
Bayesian Design → ????
Dose escalation based on probability of toxicity for the next DL

- **Response (probability of DLT)**
  - Under dosing
  - Targeted toxicity
  - Overdosing
  - Unacceptable toxicity

- **Thresholds:**
  - 20%
  - 35%
  - 60%

- **Dose escalation graph**

- **SANOFI**
How to decide the next DL to be tested?

Select the dose level with:

- Highest probability to be in the targeted toxicity interval

Safety rules:
- A Probability to be “overdosing or unacceptable tox” < 25%
- Adjacent to the tested one (No skip allowed)
Phase I trial example

Data

DLT: 0/3 11.5 mg/m²
DLT: 0/5 15.5 mg/m²
DLT: 0/3 20 mg/m²
DLT: 1/3 25 mg/m²
DLT: 0/3 25 mg/m²
DLT: 0/3 30 mg/m²

Why?
See next slide =>

Reality
Bayesian decision principle

Current dose

Estimated MTD

19.4%

17.6%
Phase I trial example

Bayesian recommendation

Data

DLT: 0/3
11.5 mg/m²

DLT: 0/5
15.5 mg/m²

DLT: 0/3
20 mg/m²

DLT: 0/3
25 mg/m²

DLT: 0/3
25 mg/m²

DLT: 0/3
30 mg/m²

DLT: 0/3
35 mg/m²

Escalate to 15.5 mg/m²

Escalate to 20 mg/m²

Escalate to 25 mg/m²

Escalate to 30 mg/m²

Escalate to 35 mg/m²

Escalate to 30 mg/m²

Escalate to 42 mg/m²

MTD
42 mg/m²

MTD
35 mg/m²

MTD
42 mg/m²

Data

DLT: 1/3
42 mg/m²

DLT: 1/3
42 mg/m²

Stay at 42 mg/m²

What is the final estimated MTD?

“3+3” → 35 mg/m²
Bayesian design → 42 mg/m²
At the end of the escalation part ...

Finally, among the 13 patients (escalation + expansion cohort) treated at 35 mg/m², 2 patients (15.4%) experienced a DLT.
For targeted anti-cancer therapies (TT), MTD may become irrelevant if therapeutic effects are already achieved at lower doses.

Worst case, the therapeutic effect may even be lower at higher doses.

Model-based phase I designs can face such a challenge:
- By finding the optimal biological dose (i.e. joint assessment of toxicity and efficacy)
- Identify a range of doses and do a randomized phase II dose-finding trial
Increasing toxicity

Over Toxic
(More than 20% ocular tox.)

20%

Useless Moderate Target
(Less than 20% ocular tox. and more than 40% resp.)

20% 40%

Increasing activity / efficacy

Toxicity vs Activity (2/2)
Balancing probability of ocular toxicities and probability of tumor response
Increasing activity / efficacy (disease resp.)

Increasing toxicity (proba ocular tox.)

Proba tox = 15%
Proba resp = 50%

Useless

Target

Moderate

Over Toxic

20%

20% 40%

Increasing activity / efficacy (disease resp.)
Balancing probability of ocular toxicities and probability of tumor response

Probability of ocular toxicity : Probability of tumor response

Plane (Probability of ocular toxicity : Probability of tumor response)
Why using the Bayesian approach?

- Bayesian design show better performances than the algorithmic « 3+3 »
- Decision tool
- Takes uncertainty into account
- Able to handle prior information when wishable
- Modeling approach: Assessment of the dose-toxicity relationship
  - Probability of toxicity is assessed whatever the dose:
    - Range of targeted toxicity can be chosen (not only 33%)
    - Ability to recommend a « better » intermediate dose (MTD between two tested dose level)
  - Allows for mechanistic based approach (takes other “endpoints” into account, e.g. PK, biomarkers …)
  - Can handle “multidrug” approaches (Combo)
I-SPY 2 clinical trial

- Adaptive screening phase II clinical trial
- Locally advanced breast cancer, neoadjuvant setting
- Primary endpoint pCR (pathologic complete response) after 5 months

Trial Objective:
- To learn as quickly as possible about efficacy of novel drugs in combo with standard chemo
- Identify treatments for patients subsets on the basis of biomarker signature
- Use earlier efficacy endpoints (MRI-based, longitudinal data)

- 5 experimental drug simultaneously

Trial adaptation
- Sample size for each experimental can vary from 20 to 120
- Experimental drugs can be dropped or graduated
- New experimental arms can come in the trial
- Bayesian adaptive randomization
Possible adaptive confirmatory clinical trials

- **Adaptive design**
  - Use accumulating data to decide on how to modify aspects of the trial without undermining the validity and integrity of the trial

- **Adaptations can include**
  - Early stopping (futility, early rejection)
  - Sample size re-assessment
  - Treatment arms (dropping, adding arms)
  - Hypotheses (Non-inferiority vs. superiority)
  - Population (inclusion/exclusion criteria; subgroups)
  - Combine trial / treatment phases

- **Bayesian tools for interim monitoring**
  - Posterior distribution of parameter of interest: repeat the hypothesis test during the course of the trial
  - Predictive probability: assess the probability that the final hypothesis test will be successful
Guidance for Industry and FDA Staff

Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

Document issued on: February 5, 2010

The draft of this document was issued on 5/23/2006
Conclusion

● More use of adaptive bayesian methods in oncology early phase clinical trials
  ● Many attractive facets for data monitoring and analysis
  ● Take into account uncertainty
  ● Prior data can help for small trials
  ● Complex data analysis models
  ● Computation easier than before

● Regulatory hurdle is high for phase III trials but … door is opening
  ● Bayesian interim analysis stopping rules
  ● Medical device FDA guidance
  ● Simulation of operating characteristics is mandatory and critical

● Perspectives
  ● Broader use of adaptive designs in oncology phase I and II clinical trials
  ● Use of more complex Bayesian modeling techniques for dose-finding trials (e.g. use of PK data, hierarchical models, mechanistic modeling)


Backup
Modèle Dose–réponse (DR)

- **Données :** N-uplets \((Y_1, \ldots, Y_N)\)
  
  où \(Y_i \sim B(n_i, \pi(d_j|\alpha_1, \beta)))\)

- **Modèle DR logistique à 2 paramètres :**

  \[
  \text{logit}(\pi(d|\alpha_1, \beta))) = \ln(\alpha_1) + \beta \ln \left( \frac{d}{d^*} \right), \quad \alpha_1, \beta > 0
  \]

  - \(d\) est la dose courante de l’agent
  - \(d^*\) est la dose de référence
  - \(\alpha_1\) et \(\beta\) sont les paramètres du modèle
  - \(\alpha_1 = \frac{\pi(d^*)}{1-\pi(d^*)}\) est l’odds (la cote) de toxicité de l’agent au niveau de dose \(d^*\)
  - Pour deux doses \(d_i\) et \(d_j\), \(\beta\) est essentiellement égal au log-odds ratio d’une DLT :

  \[
  \beta = \frac{\text{logit}(\pi(d_j)) - \text{logit}(\pi(d_i))}{\log \left( \frac{d_j}{d_i} \right)}
  \]
Toxicity vs Activity (1/2)

- Very "bad" dose
- Safe dose but not active
- Increasing activity / efficacy
- Very "good" dose
- Active dose but not safe

Increasing toxicity

Increasing activity / efficacy
Dose toxicity and dose efficacy curves

Dose-toxicity and Dose-Response curves

- DLT probability
- Credibility Interval [10%, 90%]
- Response probability
- Credibility Interval [10%, 90%]

Probability of ocular toxicity / tumor response vs. Actual dose intensity (mg/m²/week - log scale)
<table>
<thead>
<tr>
<th></th>
<th>Algorithmic (&quot;3+3&quot;)</th>
<th>Bayesian DR- model based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementatio</td>
<td>Easy</td>
<td>More complex due to statistical component</td>
</tr>
<tr>
<td>Flexibility</td>
<td>Not very flexible</td>
<td>Flexible: allows for</td>
</tr>
<tr>
<td></td>
<td>● fixed cohort size</td>
<td>● different cohort sizes</td>
</tr>
<tr>
<td></td>
<td>● fixed doses</td>
<td>● intermediate doses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Pursue several doses (schedule) in parallel</td>
</tr>
<tr>
<td>Build-up information /</td>
<td>Empirical</td>
<td>Prior information</td>
</tr>
<tr>
<td>&quot;learning process&quot;</td>
<td></td>
<td>Data gathered during the trial: DLT Can be extended to adjust for covariates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jointly model DLT and PD endpoints</td>
</tr>
<tr>
<td>Inference for true DLT</td>
<td>Observed DLT rates</td>
<td>Full inference, uncertainty assessed for true DLT rates (as dose response relationship)</td>
</tr>
<tr>
<td>rates</td>
<td>only</td>
<td></td>
</tr>
<tr>
<td>Statistical requirements</td>
<td>None</td>
<td>“reasonable” model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation required to assess behavior</td>
</tr>
</tbody>
</table>