Using Bayesian Analysis in Randomised Phase II Trials to Plan Phase III

Lucinda Billingham
Professor of Biostatistics
Director, MRC Midland Hub for Trials Methodology Research
Biostatistics Lead, Cancer Research UK Clinical Trials Unit

University of Birmingham

MRC HTMR Network Workshop, London, March 16th 2010
Using Existing Data to Inform Clinical Trial Design
Acknowledgements

• Professor Philip Johnson, University of Birmingham
• Professor Keith Abrams, University of Leicester
• Progen Pharmaceuticals
• Cancer Research UK

Reference:
David J Spiegelhalter, Keith R Abrams, Jonathan P Myles; Bayesian Approaches to Clinical Trials and Health-Care Evaluation; Wiley 2004
Agenda

• Introduction
  – Bayesian analysis
  – Hazard ratios
  – Randomised phase II trials
• Application of Bayesian analysis to randomised Phase II trials
  – Illustrative example in HCC
  – Why is it a potentially useful approach
  – How to do it
  – Interpretation of results
• Application of Bayesian analysis in seamless Phase II / III setting
• Extensions to methodology
• Objections to Bayesian methods
Aim of Statistical Analysis

Population

What is the effect of the new treatment on patient outcome compared to the standard treatment?

Sample

Data

Statistical Inference

Classical / frequentist analysis:
Estimate treatment effect with 95% confidence intervals
Statistically test hypothesis → p-value
What is a Bayesian Approach to Analysis?

• Method of statistical analysis based on theorem devised by Reverend Thomas Bayes (1702-1761)

\[ p(B / A) = \frac{p(A / B) \times p(B)}{p(A)} \]

• Alternative method to the classical / frequentist approach
  – ‘Many practising statisticians are fairly ignorant of the methods used by the rival camp and too busy to have time to find out’ Bland and Altman BMJ 1998, 317: 1151

• Acknowledges that the unknown quantity of interest is not a fixed value but could be any value with an associated probability
Bayesian Approach to Analysis

Data from trial on treatment effect

Prior probability distribution for treatment effect

Enables prior evidence or beliefs to be incorporated into the final estimate of the treatment effect

Posterior probability distribution for treatment effect

Enables direct probability statements to be made about treatment effects

Classical / Frequentist analysis just based on this
Advantages of a Bayesian Analysis

Classical

• Results are in the form of a p-value

\[ p\text{-value} = p \left( \text{data} \mid \text{no treatment effect} \right) \]

Bayesian

• Results are in the form of a probability distribution for the treatment effect

• Allows direct probability statements to be made about treatment effects

\[ \text{posterior} \rightarrow p \left( \text{treatment effect} \mid \text{data, prior} \right) \]
Measuring Treatment Effect as a Hazard Ratio (HR)

- Specific summary measure for survival data
- Measures the relative survival experience of two groups
- Hazard Ratio = \( \frac{\text{Hazard of death on New}}{\text{Hazard of death on Standard}} \)
  where the hazard is the instantaneous risk of death at any point in time
- Interpretation for survival
  \( HR = 1 \Rightarrow \) no difference between treatments
  \( HR < 1 \Rightarrow \) New treatment superior
  \( HR > 1 \Rightarrow \) New treatment inferior
- Often work with In HR as tends to have normal distribution
Phase I

What is a safe dose to give for the NEW treatment and with what toxicities?

Phase II

Is the efficacy of the NEW treatment worthy of direct comparison to STANDARD treatment of the day?

Phase III

How does the NEW treatment compare to the STANDARD treatment of the day in terms of efficacy?

Toxicities

Intermediate outcome of efficacy: Response

Overall outcome of efficacy: Survival time
Single Arm Phase II Trial

Eligible Patients

NEW Treatment

Response rate

Historical data / clinical experience of standard treatment

Problem: is the response rate better because of different patient populations?

0% Benchmark response rate 100%
Randomised Phase II Trial

Eligible Patients
Randomised

STANDARD
Response Rate

NEW1
Response Rate

NEW2
Response Rate

NEW3
Response Rate

0% Benchmark response rate

? ‘Pick the winner’ 100%
Possible Phase II / Phase III Trial Designs

Randomised Phase II

Randomised Phase III

Seamless phase II/III (e.g. Inoue, Thall, Berry; Biometrics 2002)

Randomised Phase II

Randomised Phase III

Decision Point

Should we proceed to phase III?
Current Practice for the Analysis of Randomised Phase II Trials

• Estimates and confidence intervals
  – Not clear how decision to proceed is made

• Hypothesis testing
  – Often used inappropriately so RPII just looks like underpowered PIII
  – How do the results help in decision to proceed?

Lack of knowledge on how to appropriately analyse randomised phase II trials
Goal of trial: To explore possible efficacy of PI-88 in reducing early tumour recurrence in patients who have had primary liver cancer tumours removed by surgery in order to make a decision to move to Phase 3 clinical development.

Part of a more complex design with 2 different doses and using Simon’s 2 stage study design.

Example: Randomised Phase II Trial of PI-88 in Hepatocellular Carcinoma

P-J Chen EASL 2007; Progen Pharmaceuticals

Eligible Patients with HCC who have had curative resection

Control

160 mg PI-88

Resection

4-6 weeks

Begin treatment
4 days/week
3 weeks/4 weeks
For 36 weeks or until recurrence

36 weeks

Follow up

12 weeks

Primary Endpoint: disease free rate at 48 weeks
Disease-free survival analysis
P-J Chen EASL 2007; Progen Pharmaceuticals

Dose
- Control  N=58
- 160mg  N=56

HR=0.59 (42 events) Log rank Mantel-Haenszel  p= 0.0867

~70th percentile:
Control = 27 weeks
160mg = 48 weeks

70.2% 53.9%

N=58 N=56

Should they proceed to a Phase III trial?
What Do Researchers Really Want to Know?

- Given the observed treatment effect in the randomised phase II trial (and other prior knowledge)
  - What is the likely value of the true treatment effect?
  - What is the predicted result for the planned phase III trial?
  - What are the chances of getting a statistically significant result if we continue to a phase III?

Bayesian analysis will give these answers
Bayesian Analysis in Clinical Trials

- Recommended approach for monitoring of randomised Phase III clinical trials
  - e.g. Parmar et al Lancet 2001; Berry Nature Reviews 2006
  - Aids decision-making regarding stopping a trial early
- Not explicitly been talked about for randomised phase II, but natural extension from monitoring context
### Outcome Measures: Phase II versus Phase III

<table>
<thead>
<tr>
<th></th>
<th>Phase II</th>
<th>Phase III</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td>Response rate</td>
<td>Survival time plus others</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td>Survival time</td>
<td>Response rate plus others</td>
</tr>
</tbody>
</table>

- **Outcome Measures**: Primary and Secondary measures.
- **Phase II**: Response rate, Survival time plus others.
- **Phase III**: Survival time, plus others.
- **Cross-Tabulation**: Comparison between Phase II and Phase III measures.
Bayesian Analysis

- Unknown parameter of interest is treatment effect measured in terms of log hazard ratio
  \[ \theta = \ln (\text{HR}) \]
- Bayes theorem for unknown parameter \( \theta \)

\[ p(\theta | y) \propto p(y | \theta) \times p(\theta) \]

- Posterior distribution for \( \theta \)
- Likelihood function for \( \theta \)
- Prior distribution for \( \theta \)

- Conjugate normal analysis
  - Normal likelihood so use normal prior distributions
Bayesian Analysis of PI-88 HCC Trial

Data from trial on treatment effect

Prior probability distribution for treatment effect

Posterior probability distribution for treatment effect

Aim: estimate treatment effect
i.e. Hazard Ratio (HR)

Calculations based on \( \ln HR \)

\( HR = 1 \) → \( \ln HR = 0 \)

\( HR < 1 \) → \( \ln HR \) negative

\( HR > 1 \) → \( \ln HR \) positive

Conjugate normal analysis makes calculations straightforward
Data from Trial: Likelihood Function

\[ y_m | \theta \sim N(\theta, 4/m) \quad \text{where } m = \text{number of events} \]

(Tsiatis 1981)

Trial data gave HR=0.59, m=42
N (-0.53 , 4/42=0.0952)
Prior Distributions

\[ \theta \sim N(\mu_0, 4/m_0) \quad \text{where} \quad m_0 = \text{number of events} \]

- **Non-informative**
  - $N(0, 40000)$
  - $m_0 = 0.0001$

- **Sceptic**
  - $N(0, 0.08)$
  - $m_0 = 50$

- **Plausible Enthusiast**
  - $N(-0.26, 0.04)$
  - $m_0 = 100$

- **Extreme Sceptic**
  - $N(0.26, 0.04)$
  - $m_0 = 100$
Posterior Distributions (1)

\[ \theta | y_m \sim N \left( \frac{m_0 \mu_0 + my_m}{m_0 + m}, \frac{4}{m_0 + m} \right) \]

Non-informative

Plausible Enthusiast

- **PI-88 Superior**
  - \( p(\text{HR}<1) = p(\ln \text{HR}<0) = 0.96 \)
  - \( p(\text{HR}<0.75) = p(\ln \text{HR}<-0.29) = 0.78 \)

- **Control Superior**
  - \( p(\text{HR}<1) = p(\ln \text{HR}<0) = 0.98 \)
  - \( p(\text{HR}<0.75) = p(\ln \text{HR}<-0.29) = 0.62 \)
Posterior Distributions (2)

Sceptic

- $p(HR<1) = p(\ln HR<0) = 0.88$
- $p(HR<0.75) = p(\ln HR<-0.29) = 0.41$

Extreme Sceptic

- $p(HR<1) = p(\ln HR<0) = 0.44$
- $p(HR<0.75) = p(\ln HR<-0.29) = 0.03$
### Summary of Posterior Results

<table>
<thead>
<tr>
<th></th>
<th>Posterior</th>
<th>P(HR&lt;1)</th>
<th>P(HR&lt;0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-informative</strong></td>
<td>N(-0.53,0.0952)</td>
<td>0.96</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Plausible Enthusiast</strong></td>
<td>N(-0.34, 0.0282)</td>
<td>0.98</td>
<td>0.62</td>
</tr>
<tr>
<td><strong>Sceptic</strong></td>
<td>N(-0.24, 0.0435)</td>
<td>0.88</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>Extreme Sceptic</strong></td>
<td>N(0.026, 0.0282)</td>
<td>0.44</td>
<td>0.03</td>
</tr>
</tbody>
</table>
Predictive Distributions (1)

\[ Y_n \mid y_m \sim N \left( \frac{m_0 \mu_0 + m y_m}{m_0 + m}, 4 \left( \frac{1}{m_0 + m} + \frac{1}{n} \right) \right) \]

Plan new trial with 300 events; increase variance of posterior by \( \frac{4}{300} = 0.0133 \)

Non-informative Plausible Enthusiast

**PI-88 Superior**

\[ p(HR<1) = p(\ln HR<0) = 0.95 \]
\[ p(HR<0.75) = p(\ln HR<-0.29) = 0.77 \]

**Control Superior**

\[ p(HR<1) = p(\ln HR<0) = 0.95 \]
\[ p(HR<0.75) = p(\ln HR<-0.29) = 0.60 \]
Predictive Distributions (2)

Sceptic

P(HR<1) = p(lnHR<0) = 0.84
P(HR<0.75) = p(lnHR<-0.29) = 0.42

Extreme Sceptic

P(HR<1) = p(lnHR<0) = 0.45
P(HR<0.75) = p(lnHR<-0.29) = 0.06
## Summary of Predictive Results

<table>
<thead>
<tr>
<th></th>
<th>Posterior Predictive</th>
<th>P(HR&lt;1)</th>
<th>P(HR&lt;0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-informative</td>
<td>N(-0.53, 0.0952)</td>
<td>0.96</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>N(-0.53, 0.1086)</td>
<td>0.95</td>
<td>0.77</td>
</tr>
<tr>
<td>Plausible Enthusiast</td>
<td>N(-0.34, 0.0282)</td>
<td>0.98</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>N(-0.34, 0.0415)</td>
<td>0.95</td>
<td>0.60</td>
</tr>
<tr>
<td>Sceptic</td>
<td>N(-0.24, 0.0435)</td>
<td>0.88</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>N(-0.24, 0.0568)</td>
<td>0.84</td>
<td>0.42</td>
</tr>
<tr>
<td>Extreme sceptic</td>
<td>N(0.026, 0.0282)</td>
<td>0.44</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>N(0.026, 0.0415)</td>
<td>0.45</td>
<td>0.06</td>
</tr>
</tbody>
</table>
Hybrid Classical-Bayesian Approach to Power

- Assume conclusions of trial will be based entirely on classical analysis
- Classical power = p( reject H0 | \( \theta = \theta^* \) )
- Use predictive distribution to calculate the overall unconditional probability of a ‘classically’ significant result
  - ‘Expected power’
Predictive Probability of Obtaining a ‘Classically’ Significant Result in New Trial

n=300, significance level = 5%
Classical power = \( p(\text{reject } H_0 \mid \theta^*=-0.29 \text{ ie } HR^*=0.75) = 0.70 \)

\[
Y_n \mid y_m \sim N\left( \frac{m_0 \mu_0 + my_m}{m_0 + m}, \sqrt{\frac{1}{m_0 + m} + \frac{1}{n}} \right)
\]

\[
\text{Power}_C = \Phi\left[ \sqrt{\frac{m_0 + m}{m_0 + m + n}} \left( \frac{\mu_n \sqrt{n}}{2} + z_\epsilon \right) \right]
\]

<table>
<thead>
<tr>
<th></th>
<th>Power (n=300)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-informative</td>
<td>82%</td>
</tr>
<tr>
<td>Plausible Enthusiast</td>
<td>71%</td>
</tr>
<tr>
<td>Sceptic</td>
<td>53%</td>
</tr>
<tr>
<td>Extreme Sceptic</td>
<td>11%</td>
</tr>
</tbody>
</table>
Hybrid Classical-Bayesian Power Curves
‘Bayesian Power’

- Assume conclusions of trial will be based on Bayesian analysis
- Define Bayesian significance
  \[ p(\theta > 0|\text{data}) < \varepsilon \]
- Use predictive distribution to calculate the expected ‘Bayesian’ power, averaged with respect to the prior distribution
Predictive Probability of Obtaining a ‘Bayesian’ Significant Result in New Trial

\[ Y_n | y_m \sim N \left( \frac{m_0 \mu_0 + m y_m}{m_0 + m}, 4 \left( \frac{1}{m_0 + m} + \frac{1}{n} \right) \right) \]

\[ \text{Power}_B = \Phi \left[ \frac{\mu_n \sqrt{m_0 + m + n} \sqrt{m_0 + m}}{2 \sqrt{n}} + \sqrt{\frac{m_0 + m}{n}} z_{\epsilon} \right] \]

<table>
<thead>
<tr>
<th></th>
<th>Power (n=300)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-informative</td>
<td>86%</td>
</tr>
<tr>
<td>Plausible Enthusiast</td>
<td>87%</td>
</tr>
<tr>
<td>Sceptic</td>
<td>60%</td>
</tr>
<tr>
<td>Extreme Sceptic</td>
<td>6%</td>
</tr>
</tbody>
</table>
Bayesian Power Curves
Example: Phase II/III
Inoue, Thall & Berry Biometrics 2002

• NSCLC trial, E vs S, n=900, 72 months
• \( \phi(t) = p(\Delta > 0 \mid D_{72}) \)
• Large \( \phi(t) \) \( \Rightarrow \) if maximum allowed future resources were expended then likely that E>S
• Decision based on predictive probabilities involving future data at 72 months
• PII to PIII decision: analysis at t=8, 10, 12 months
  - \( 0.01 < P(\phi(t) > 0.98) < 0.80 \) then continue PII
  - \( P(\phi(t) > 0.98) \geq 0.80 \) then organise PIII
  - \( P(\phi(12) > 0.98) < 0.80 \) then conclude E<S
Extensions to Methodology

• Consider other priors: lump and smear, evidence-based
• Response rate as primary outcome measure
  – Binomial likelihood
  – Beta prior
  – Beta-Binomial conjugate analysis
• Non-conjugate analysis
  – Use software to simulate posterior and predictive distribution
• Predicting phase III primary outcome (e.g. survival) from phase II primary outcome (e.g. response)
• Extension to include utilities (Bayesian decision theoretic approach) and costs (value of information) in the decision making
• Trial design appropriate to planned analysis
Why Do People Object to the Use of Bayesian Methods?

- Use of priors introduces an element of subjectivity
- Which priors to use
- No single measure of statistical significance
- Fear of acceptance in terms of publication and regulatory bodies
- Computational aspects
- Lack of experience and understanding
Conclusions

• Use of randomised phase II trials is increasing
• No clear guidance on how to analyse randomised phase II trials
• Bayesian analysis is promoted as method for interim analysis of phase III
• Bayesian analysis seems to be the natural approach for randomised phase II trials that will give researchers the answers they want and should be promoted