Learning as we go: the use of Bayesian Methodology for Adaptive Designs

Alun Bedding

Director

Biostatistics and Programming
Development Partners

Outline

- Introduction to Bayesian statistics and decision theory
- When to use Bayesian methods
- Continual reassessment method
- Example using safety and efficacy binary data
- Example using safety and efficacy continuous data
- Conclusions

“We are all inherently Bayesian but it is difficult to behave like one”

– Thomas Louis, Johns Hopkins Bloomberg SPH

GlaxoSmithKline

Bayesian Statistics

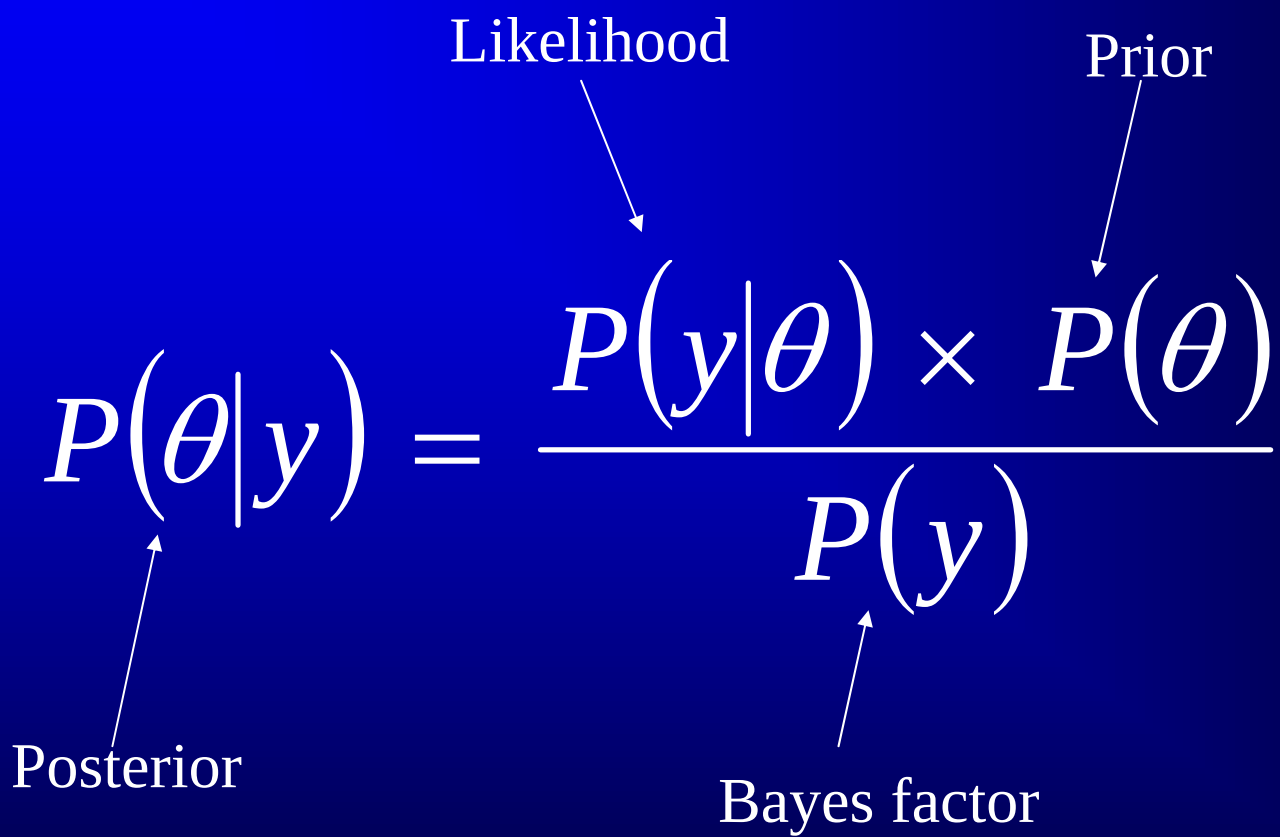
- Use data to modify prior subjective beliefs about a drug's properties into more reliable posterior beliefs.

PRIOR ---- BAYES THEOREM ----- POSTERIOR

Knew this before ---- Saw this ----- Now know this

- Assume all unknown parameters follow a probability distribution

Bayes Theorem



GlaxoSmithKline

Bayes Inference

- Bayes Theorem

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}$$

- The result is a probability distribution for θ
  - Report mean, median, mode, standard deviation
  - Credible intervals
- Probability statements $\Pr(\theta > x|y) = \alpha\%$
  - These are direct probabilities and are not available in frequentist inference

Why Bayes?

Bayesian approach has been described as:

“the explicit quantitative use of external evidence in the design, monitoring, analysis, interpretation and reporting of a health-care evaluation”

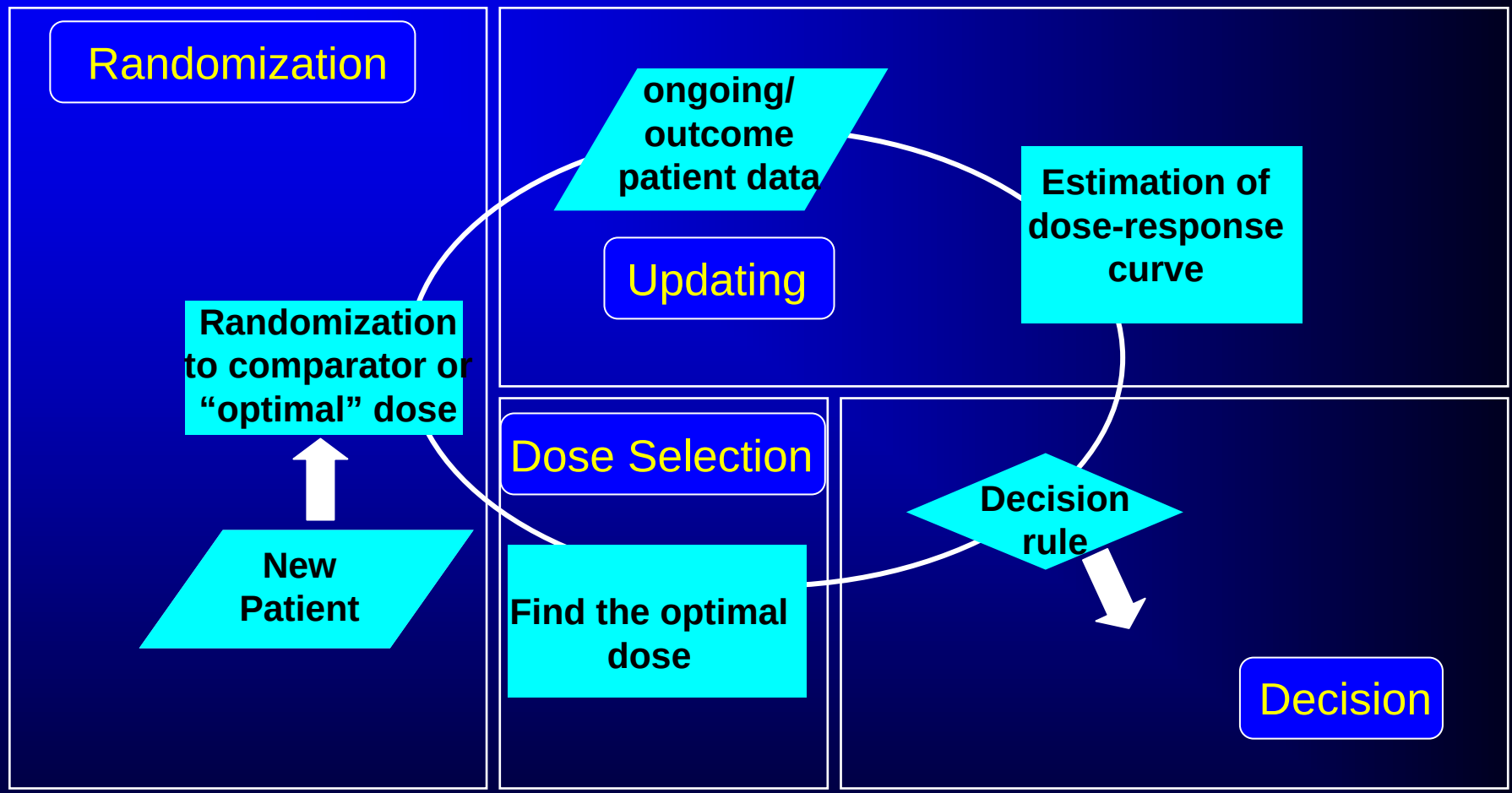
(Spiegelhalter, Abrams and Myles, 2002)

Goals

- Learn faster: More efficient trials
- More efficient drug/device development
- Better treatment of patients in clinical trials – more ethical

GlaxoSmithKline

Diagram of an Adaptive design



GlaxoSmithKline

There are Potentially Two Decision Problems

- Dose allocation – which dose to assign to next patient to learn the most about the dose-response curve
- Stopping rules
  - Stop and abandon the drug
  - Drop a dose
  - Add a dose
  - Stop and go confirmatory
  - Continue

Apply Bayesian Decision Theory

- In any given period, decision to be made as to which doses should be administered to the next patient
- First patient - dose based on prior information
- Response observed - information updated to obtain posterior - optimal choice for next patient.
- Need a utility or gain function to determine the optimal dose

Utility

- Value of a decision taken within a study
- Could be cost – what are the cost implications of making a decision
- Efficacy/safety issues – if I increase dose what happens to both
- Estimating the model
- Gain functions are very similar
- Berger – Statistical Decision Theory and Bayesian Analysis

Utilities – Simple Case

- Utility theory – Simple Case
- Assume we have 2 actions a_1 and a_2, and that the rewards for a_1 and a_2 are r_1 and r_2 respectively.
- If we now say our utilities for the rewards are $U(r_1)$ and $U(r_2)$ and the probabilities are π_1 and π_2 then the expected utility E(U(r)) is:
- $\pi_1 U(r_1) + \pi_2 U(r_2)$
- This is very simplistic!!!!!! There are other forms of utility.

When to use Bayesian Methods

- Most exploratory studies where the interest is learning about the compound
- Where you can incorporate prior information
- To ***answer*** the right ***question***
- Dose response and the decision for Phase III lends nicely to Bayesian methods
- Establish primary endpoint that follows through phases
- Monitoring safety

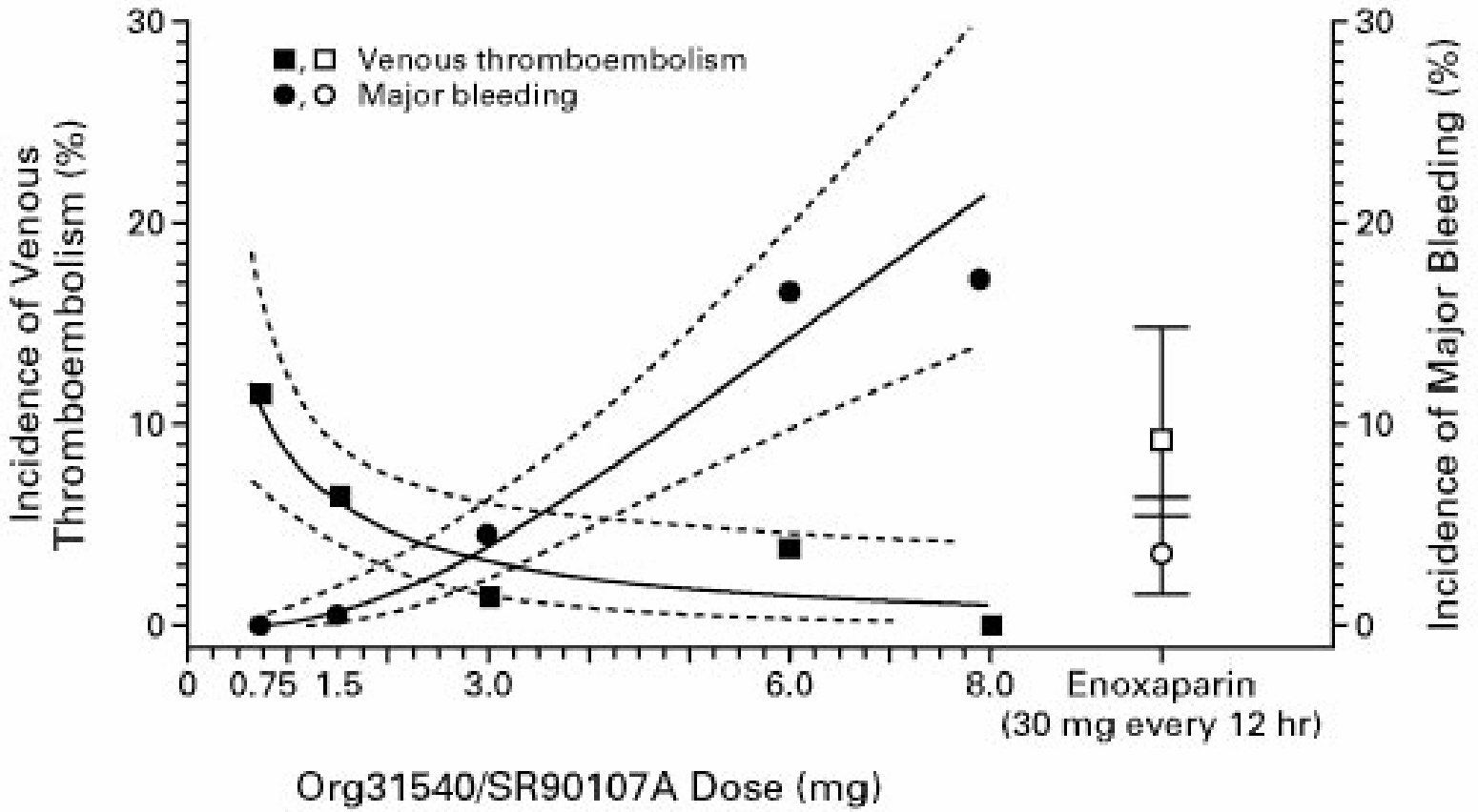
Continual Re-Assessment Method

- Identify a dose with target toxicty (i.e. TD20 – dose which gives 20% toxicity)
- Doses are pre-defined / Outcome is binary
- Uses a simple dose-response model
- Picks next dose based on the expected response to that dose is close to the TD20
- O'Quigley's own modification (1996) - start at the lowest dose and escalate 1 dose at a time, until toxicity observed

Example – New Anti-Thrombolytic Treatment

- Phase IIa study
- **Objective**
- **To identify at least one dose of Experimental Drug (ED) that is both *efficacious* in terms of venous thromboembolism (VTE) and *safe* in terms of major bleeding relative to active comparator (AC).**
  - Definitions
    - Define *efficacious* as a VTE event rate with ED that is 3 percentage points (pp) less than that achieved with AC
    - Define *safe* as a major bleeding rate with ED that is 4 pp less than that achieved with AC

Example – New Anti-Thrombolytic Treatment



Models for the Data

- Assume a bivariate logistic model to model both efficacy and safety together – call this df(θ)
- Model introduced by McCullagh and Nelder
- Bayesian approach by Cengiz (2005) Bayesian inference for bivariate generalized linear models in diagnosing renal arterial obstruction - Statistical Methodology 2 (2005) 168–174

Objective re-written

- Let π_{EE}- probability of VTE on ED
- π_{EA}- probability of VTE on AC
- π_{SE}- probability of major bleed on ED
- π_{SA}- probability of major bleed on AC
- Difference in proportions for efficacy (VTE)
  $d_e = \pi_{EE} - \pi_{EA}$
- Difference in proportions for safety (major bleed)
  $d_s = \pi_{SE} - \pi_{SA}$
- Let Z* be the dose that gives VTE rate 3 pp less and major bleed 4 pp less than AC.

Z^* = Dose for $d_e = 3$ and $d_s = 4$

GlaxoSmithKline

Choosing next dose…

- Estimate Z* as precisely as possible – minimize the posterior variance for Z* (maximise the utility around a negative posterior variance)
- **Gain function**
  - **$G(z,y_p,D) = -Var[z^*|df(\theta),y_p,z]$**
- z – assigned future dose
  y_p – predicted responses of a new patient given a dose z
  $df(\theta)$ – model for the data
- Pick dose which maximises this gain function

GlaxoSmithKline

Stopping Rules

- Study had 7 doses
- Define:
- Effective drug efficacy $E_E - d_e < 3$ pp
- Effective drug safety $E_s - d_s < 4$ pp
- Ineffective drug efficacy $I_E - d_e > 15$ pp
- Ineffective drug safety $I_s - d_s > 6$ pp

Look at Posterior Predictive Probabilities for Doses

- Joint posterior predictive distribution for safety and efficacy
- If $Pr(E_E \text{ and } E_S) > 0.9$ then stop and choose that dose for confirmatory.
- If $Pr(I_E)$ or $Pr(I_S) > 0.9$ for a dose then stop that dose.
- If $Pr(I_E)$ or $Pr(I_S) > 0.9$ for all doses then stop development.
- Else continue.

Simulations

- Simulate the design
  - Under the null hypothesis – ED=AC for both efficacy and safety - type I error
  - Under the alternative – ED-AC < 3pp for efficacy and < 4pp for safety – “Power” – how often do we make the correct positive decision
- Results assuming total of 500 patients
  - Power approx 85%
  - Type I error less than 1%

Example using continuous safety and efficacy

- ADHD is a central nervous system disorder which has its onset in childhood and has been found to occur in 3% to 5% of school-age children. ADHD frequently persists into adult life.
- Objective – reduce ADHD rating scale
- ADHD rating scale score decreases with decreasing dose – lower the score the better
- Normative ADHD score = 24 (L_E)
- Safety concern of an increase in heart rate over baseline
- Problem change in heart rate = 10 bpm (L_S)
- Have 10 doses (5, 15, 25, 35, 45, 55, 65, 75, 85, 95)
- Patients receive more than one dose

Bayesian Model for Phase II – Multivariate Data

- Let
- Y_{Eij} be the efficacy response for subject i and occasion j
- Y_{Sij} be the safety response for subject i and occasion j
- d_j is the dose given at occasion j
- Then we can assume a multivariate model:

Bayesian Model for Phase II – Efficacy and Safety

- $(Y_{Eij}, Y_{Sij})' \sim MVN[(M_{Eij}, M_{Sij})', \Omega]$
- where:
- $M_{Eij} = \theta_{E1} + \theta_{E2} d_j + S_{Ei}$
- $M_{Sij} = \theta_{S1} + \theta_{S2} d_j + S_{Si}$
- $(S_{Eij}, S_{Sij})' \sim MVN[(0, 0)', \Sigma]$
- $\Omega \sim$ Wishart(R,df)
- $\Sigma \sim$ Wishart(X,df)

GlaxoSmithKline

Gain function

- Need to gain on both efficacy and safety within each patient
- From literature review assume linear models for the responses – the ADHD on a log scale
- Let d_e = reduction in ADHD
  Let d_s = increase in heart rate
- Let Z* be the dose that gives $d_e < 24$ and $d_s < 10$ bpm

Choosing next dose for a patient or new

- Give next dose which gives best posterior predictive probability of Z* given the current posterior estimates of the unknown parameters
- **Gain function**
  - **$G(z,y_p,D) = Pr(Z^* | df(\theta),y_p,z)$**
- z – assigned future dose
  y_p – predicted responses of a new patient given a dose z
  $df(\theta)$ – model for the data
- Pick dose which maximises this gain function
- Can choose best dose combination to maximise gain in Phase III

GlaxoSmithKline

Priors

- Assume an informative prior around the model parameters – taken from Phase I results
- Using vague priors leads to odd dose schemes if you do not assume dose escalation
- Priors taken from Phase I are less subjective
  - Based on data using the same drug
- More subjective way to use priors taken from published literature

Results from a Bayesian Analysis of Phase II – Revisiting Old Phase II Data

- The following are results from a Bayesian of data from combining the data from two phase II studies for a drug to treat ADHD – combined sample size of 2750 (approx)
- New Phase II designed with 500 patients in an adaptive setting
- Informative priors assumed from Phase I
- Posterior estimates:
- $\theta_{E1} = 4.445$, $\theta_{E2} = -0.3542$, $\theta_{S1} = -1.758$, $\theta_{S2} = 3.006$
- $Pr(d_e<24, d_s<10$ bpm after 4 doses) = 0.801
- Best dose scheme = 5, 15, 25, 35

GlaxoSmithKline

Conclusions

- Bayesian decision theory methods provide a framework by which better understand data – learning framework
- Can be easily used in dose-response and dose-escalation studies
- Need real time turnaround of data and a good model for the data
- Makes dose response decisions easier
- Better probability of success in Phase III