Imperial College London

‘Bayesian Statistics: are theoretical advances changing practice in the pharmaceutical industry and regulation of medicine?’

Professor Deborah Ashby
Imperial Clinical Trials Unit
School of Public Health
Imperial College London

Outline

- Bayes- a brief history and pictorial guide
- Example: Statins
- Dose finding in early phase studies
- Risk-benefit decision making

Outline

- Bayes- a brief history and pictorial guide
- Example: Statins
- Dose finding in early phase studies
- Risk-benefit decision making


ROYAL TUNBRIDGE WELLS
THOMAS BAYES
1702 - 1761
Nonconformist minister
and mathematician
Originator of the statistical
theory of probability, the basis
of most market research and
opinion poll techniques
lived here
1731 - 1761
FOURTH CENTENARY

Imperial College
London

Imperial College
London

1763 – Thomas Bayes

An Essay Toward Solving a Problem in the Doctrine of Chances
Bayes, T. Rev. (1763)
Philos. Trans. R. Soc. London, 53, 370-418

Imperial College London

Bayes theorem

$$\text{pr}(B_j|A) = \frac{\text{pr}(A|B_j)\ \text{pr}(B_j)}{\text{pr}(A)}$$

or

$$\text{pr}(B_j|A) \propto \text{pr}(A|B_j)\ \text{pr}(B_j)$$

Note: A represents data

B_1 or B_2…. alternative explanation or hypotheses

Cox & Hinckley, 1974, chapter 10

$$f_{\theta|Y}(\theta|y) \propto f_{Y|\theta}(y|\theta)\, f_{\theta}(\theta)$$

θ and Y random variables

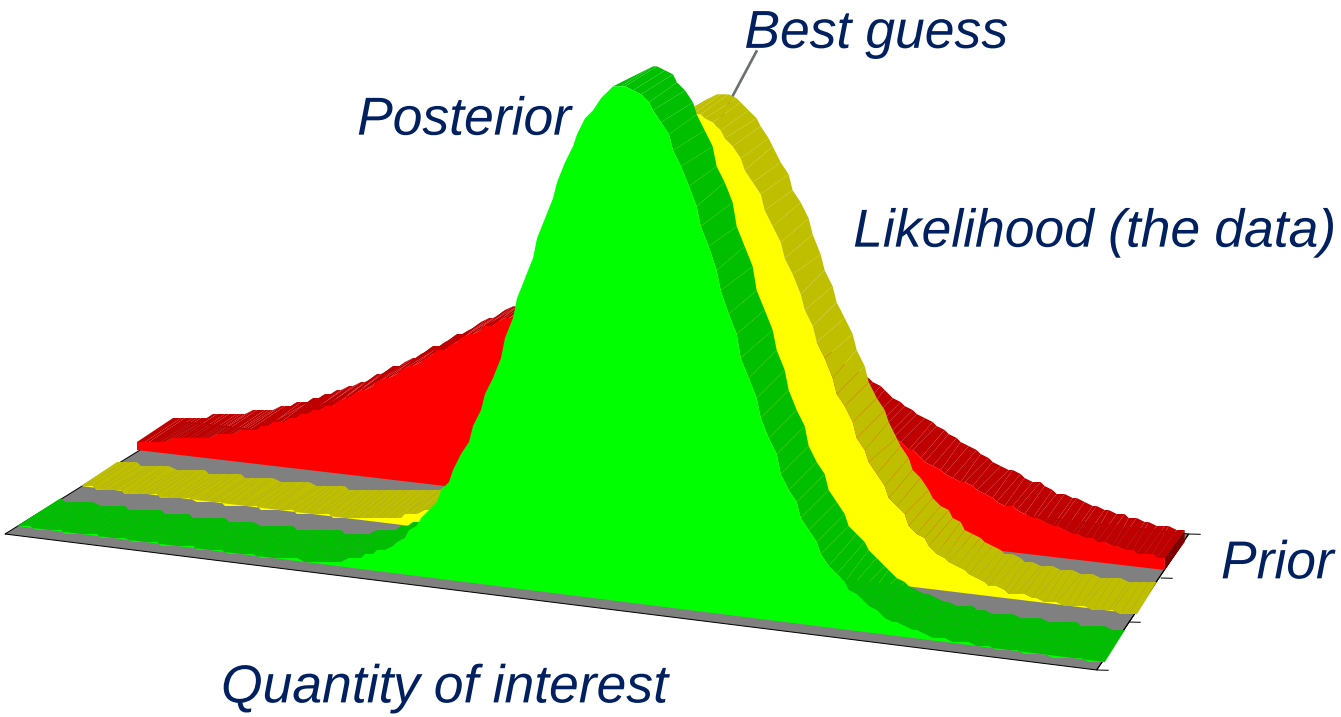
$f_{y|\theta}(y|\theta)$ likelihood, written in conditional form

- tractable using conjugate distributions
- prior distributions
  - frequency distributions
  - normative and objective representations of beliefs
  - subjective measure of individual belief
- estimation + hypothesis testing possible
- large samples mean variance depends on likelihood not prior

Cox & Hinckley, 1974, chapter 10

Imperial College
London

Bayesian approach



Outline

- Bayes- a brief history and pictorial guide
- Example: Statins
- Dose finding in early phase studies
- Risk-benefit decision making

Imperial College London

4S Trial (Scandinavian Simvastatin Survival Study)

Number of deaths/subjects			Rate ratio (95% CI)
	Placebo	Simvastatin	
Men	231/1803	155/1814	0.66 (0.53-0.80)
Women	25/420	27/407	1.12 (0.65-1.93)
All	256/2223	182/2221	0.70 (0.58-0.85)

4S Trial

Should women be treated?

No - overall effect ‘significant’;
interaction ‘non-significant.

Yes - estimated effect for women ‘non-significant’;
in wrong direction.

Imperial College
London

4S Trial

```
model
{for (i in 1:2)       # i indexes sex 0= men, 1=women
{for (j in 1:2)       # j indexes treatment,
                            0=placebo, 1=simvastatin

{logit(p[I,j])<alpha+beta.sex*sex[i]+beta.trt*trt[j]+
beta.interaction*sex[i]*trt[j]

deaths[i,j]~dbin(p[i,j],n[i,j])
}}
```

4S Trial

#priors

alpha ~ dnorm (0,0.00001)

beta.sex ~ dnorm(0,0.00001)

beta.trt ~ dnorm (0,0.00001)

beta.interaction ~ dnorm(0,0.00001)

4S Trial

or.men <- exp. (beta.trt)

or.women <- exp (beta.trt+beta.interaction)

4S Trial – non-informative priors

	Median	2.5%	97.5%
beta.trt	-0.45	-0.65	-0.23
beta.interaction	0.56	-0.06	1.15
or.men	0.64	0.52	0.79
or.women	1.12	0.63	1.96

Imperial College
London

4S Trial

informative trial on interaction

ratio of true treatment effects

95% range (0.8, 1.25)

#prior

beta interaction ~ dnorm (0, 77.2)

4S Trial- informative prior on I/A

	Median	2.5%	97.5%
or.men	0.68	0.55	0.83
or.women	0.72	0.54	0.96

Outline

- Bayes- a brief history and pictorial guide
- Example: Statins
- Dose finding in early phase studies
- Risk-benefit decision making

(Some) Current Approaches to Designing Phase 2

Select subset of doses from Phase 1

- Informal estimate of effects in patients
- Design Phase 2

Placebo/Comparator plus 3 or more doses

- Hypothesis testing → 1 or 2 doses into Confirmatory trials
- Curve-fitting (e.g. logistic) → estimation of quantities such as ED50

(Brief) Background on Dose-Finding

ICH E4 “Dose-Response Information to Support Drug Registration”

- “Knowledge of the relationships among dose, drug concentration in blood, and clinical response (effectiveness and undesirable effects) is important for the safe and effective use of drugs in individual patients.”

Propose dose(s) for larger trials (Phase 2b or 3)

Identify therapeutic window

Minimum Effective Dose (MED) to Maximum Tolerated Dose (MTD)

Life-threatening (e.g. oncology) versus other diseases

Identify frequency and duration of dosing

Further Guidance from ICH E4

*"What is most helpful in choosing the starting dose of a drug is **knowing the shape and location of the population (group) average dose-response curve** for both desirable and undesirable effects. Selection of dose is best based on that information,…"*

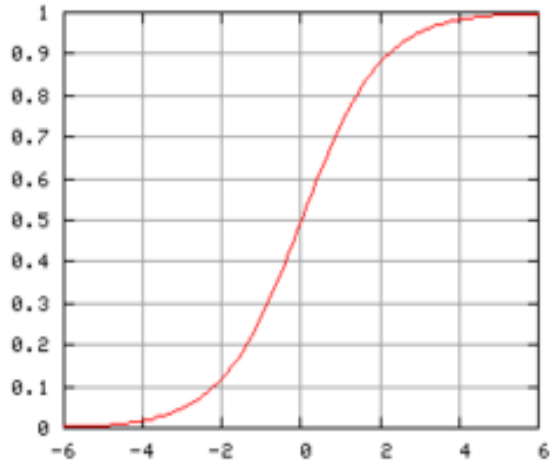
*"In principle, being able to detect a **statistically significant difference in pair-wise comparisons between doses is not necessary if a statistically significant trend (upward slope) across doses can be established** using all the data. It should be **demonstrated, however, that the lowest dose(s) tested, if it is to be recommended, has a statistically significant and clinically meaningful effect**"*

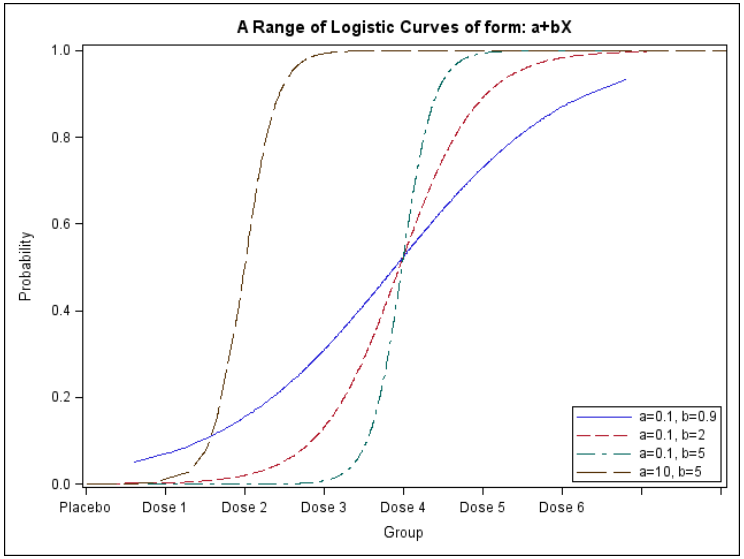
*"Study designs usually should emphasize **elucidation of the dose-response function, not individual pair-wise comparisons**. If a particular point on the curve, e.g., whether a certain low dose is useful, becomes an issue, it should be studied separately."*

Imperial College London

Fitting a Dose-Response Curve

Logistic distribution





- *Can a model provide more information?*
  - *On doses not studied?*
  - *Estimation (vs hypothesis testing)*
  - *Assign/recruit more patients to steep section*
- *Potential drawbacks*
  - *Assumes monotonically increasing*
  - *Harder to summarise to non-statisticians*
  - *Bayesian: need to elicit Priors for parameters*

$$p(y=1)=\frac{e^{a+bx}}{1+e^{a+bx}}=\frac{1}{1+e^{-(a+bx)}}$$

- *Variations*
  - *Not just for modeling Probabilities*
  - *"Shapes" other than logistic too*

Some Reasons to Improve

Sub-optimal current status:

- High Failure rates in Phase 3
  - » 45% (Accenture, 2001)
- High frequency of post-approval amendments to dosing information in label
  - » E.g. FDA Post-Marketing Commitments review (PDUFA):
    - 245 products with 743 PMCs (2002-2005)
    - 51% of fulfilled PMCs resulted in label change
    - Most common reason “safety & efficacy concerns” (30%)
- What if Phase 2 was a “fluke” (i.e. outlier from expectations)
  - » Unrealistic assumptions for Phase 3

Some Reasons to Improve (2)

Opportunities

- Wider acceptance & software to apply methods to (formally) quantify prior knowledge (Bayesian)
- Fewer NCEs but increase in line extensions & new populations
- Translational Medicine, Disease models and pharmacodynamic endpoints/biomarkers quantify knowledge

→ **More opportunity to apply Bayesian methods in a "learning environment"**

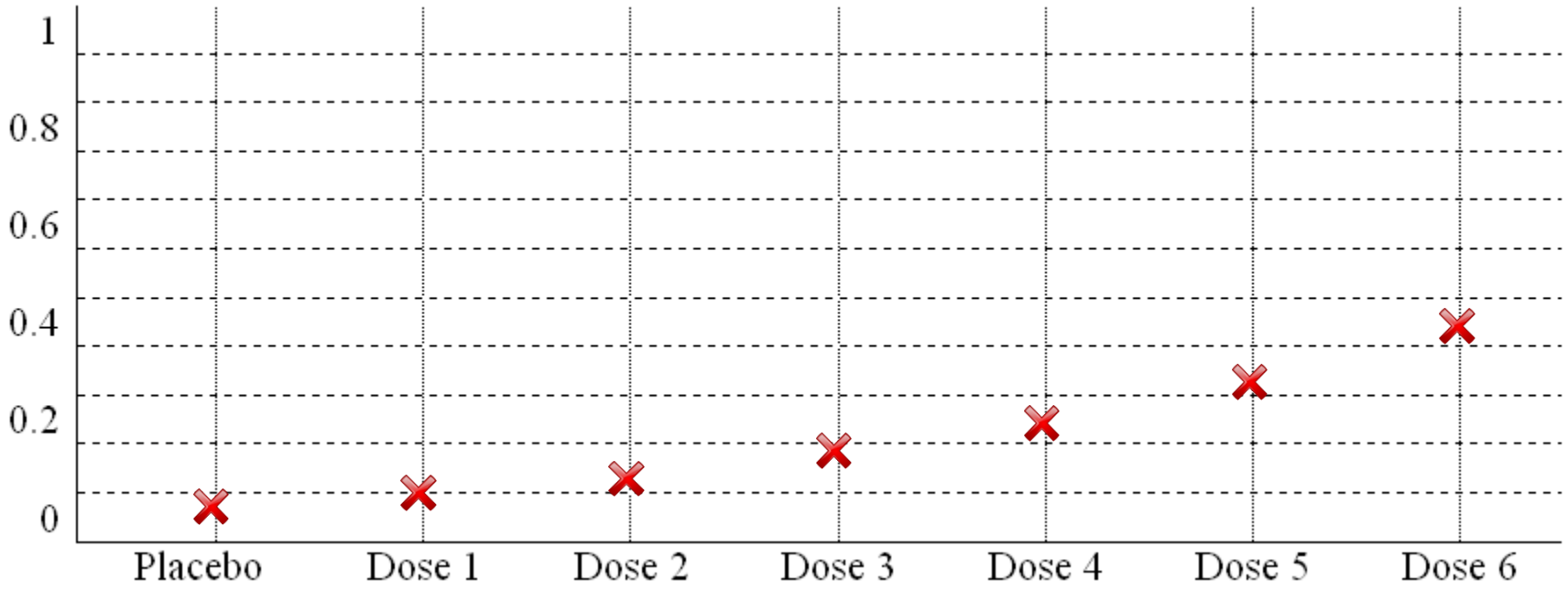
Recent Work in this Area

Huson & Kinnersley (2009)

- 1 Stage Elicitation
- Incorporate experts' beliefs (indirectly)
- Form Priors for (logistic) parameters
- Case Study
  - » Placebo + 6 doses
  - » Probability of specific toxicity (n=5 experts)
- General framework

Imperial College London

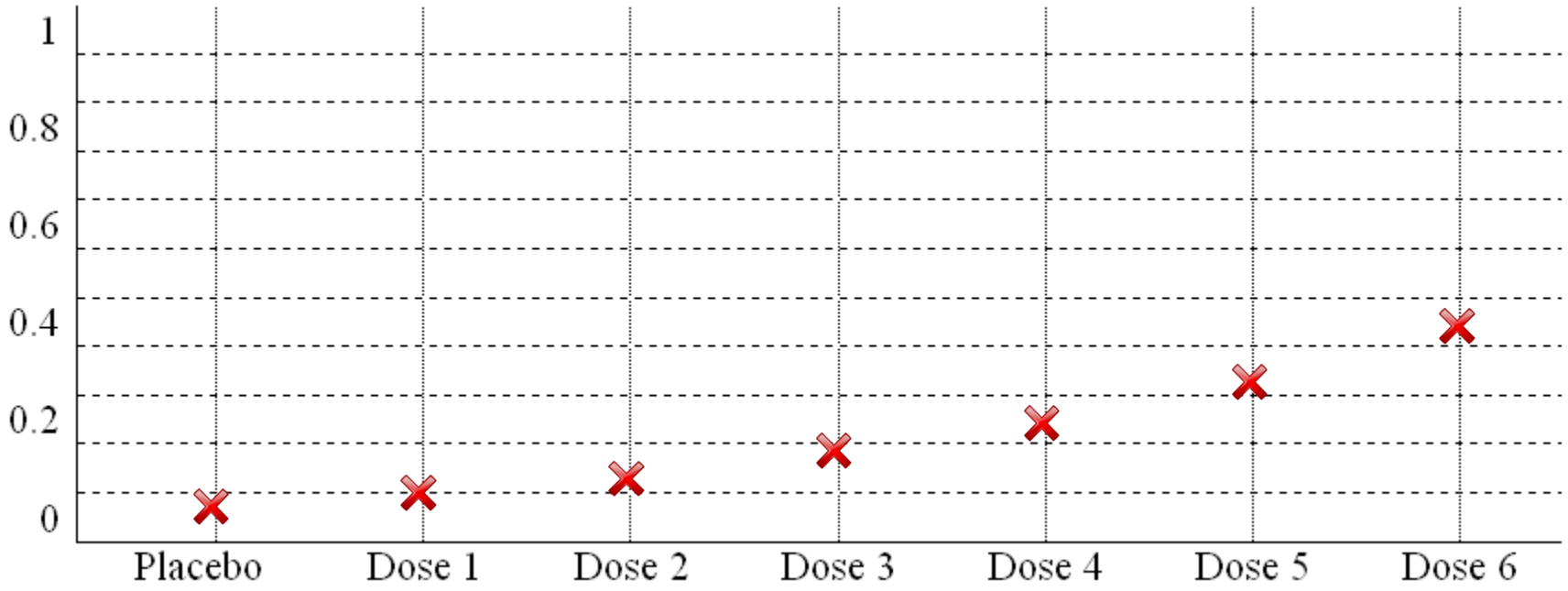
Case Study of Huson & Kinnersley



What are the likely values of specified toxicity for each dose group?

Imperial College London

Case Study of Huson & Kinnersley



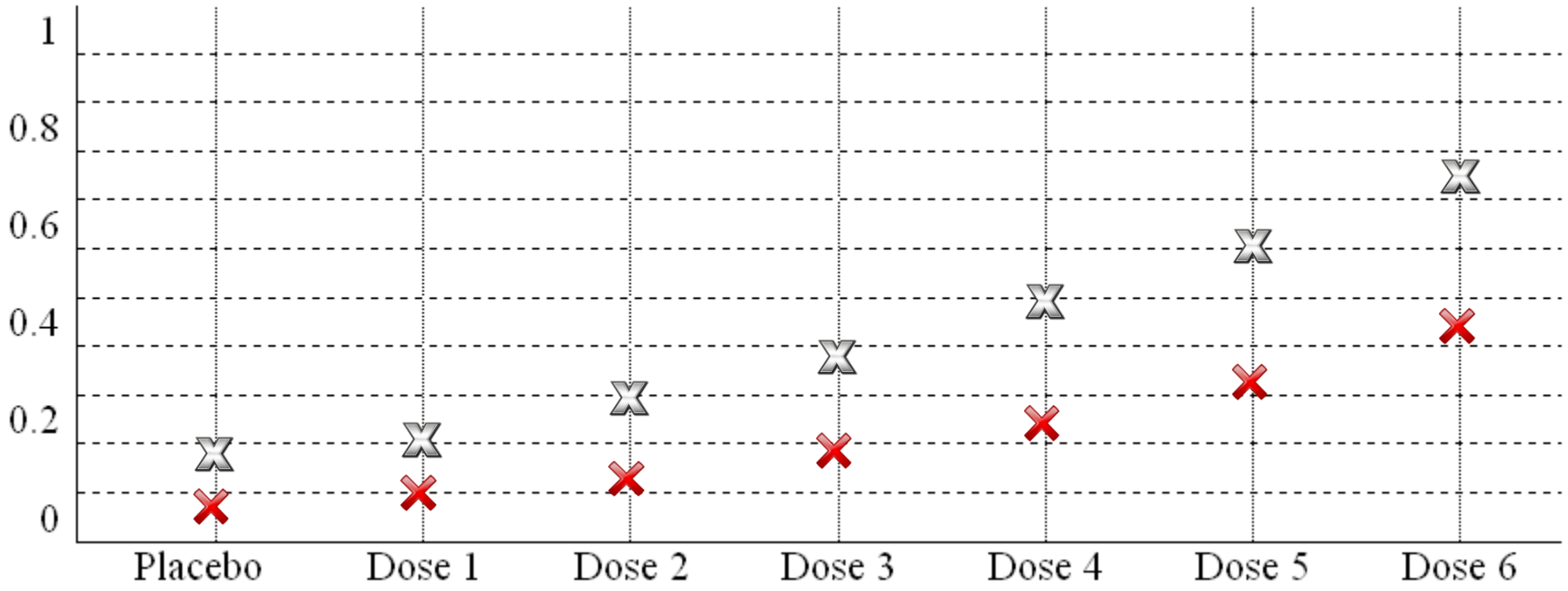
Suppose this expert's belief can be described by a logistic model

$1 / \{1 + e^{-(a+bx)}\}$ *with* $a=-4.8$ *and* $b=0.44$

But for Bayesian approach, we need a distribution for "a" and "b"

Imperial College London

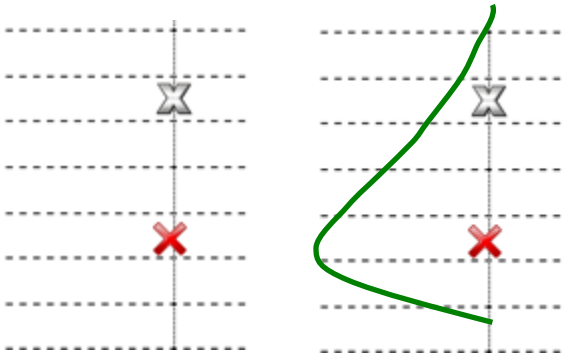
Case Study of Huson & Kinnersley

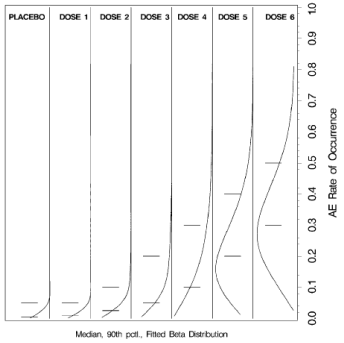


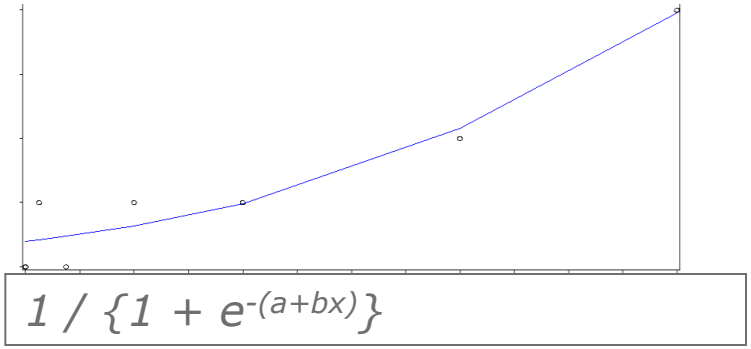
Acknowledge variability within an expert *→ simulate a range of values for each expert (at each dose)*
Provide estimates which you think are highly unlikely to occur *i.e. 90% sure the response will not occur?*

Imperial College London

Case Study of Huson & Kinnersley





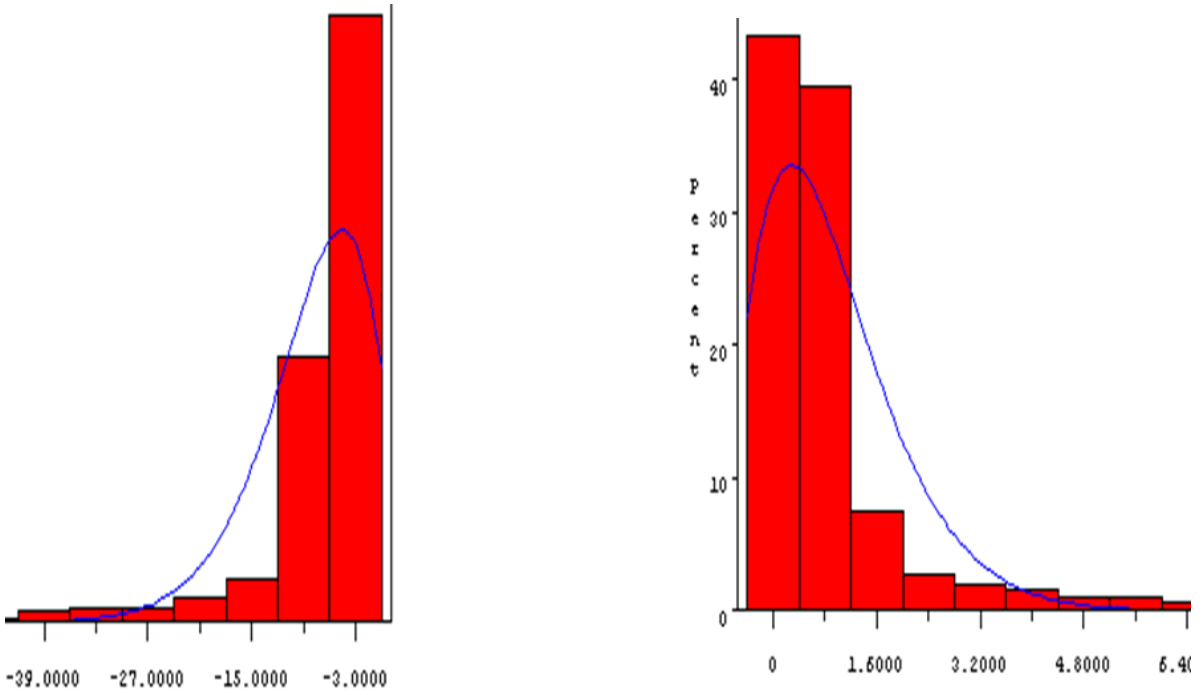


For each dose:

- *Impose distribution (e.g. beta)*
- *Beta(α,β) describes green curve*
- *Simulate values from that beta distribution*
- *Fit logistic curve across all doses*
- *Store "a" and "b" then repeat 1000s times*

Imperial College London

Case Study of Huson & Kinnersley



Empirical Distributions for "a" and "b" (Intercept and Dose Effect)

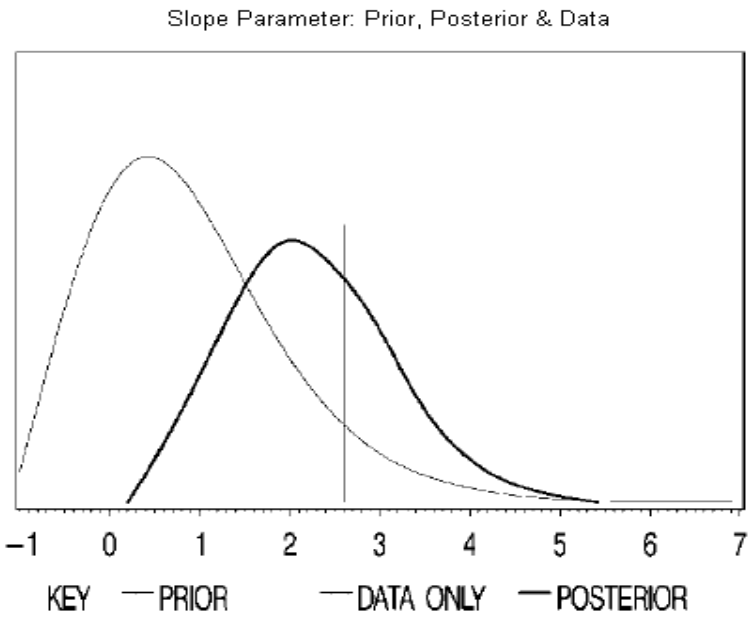
Blue curves now represent our Priors for "a" and "b" (we chose beta)

Imperial College London

Case Study of Huson & Kinnersley

Run trial

→ *Combine Prior with Study Data to form Posterior Distribution*

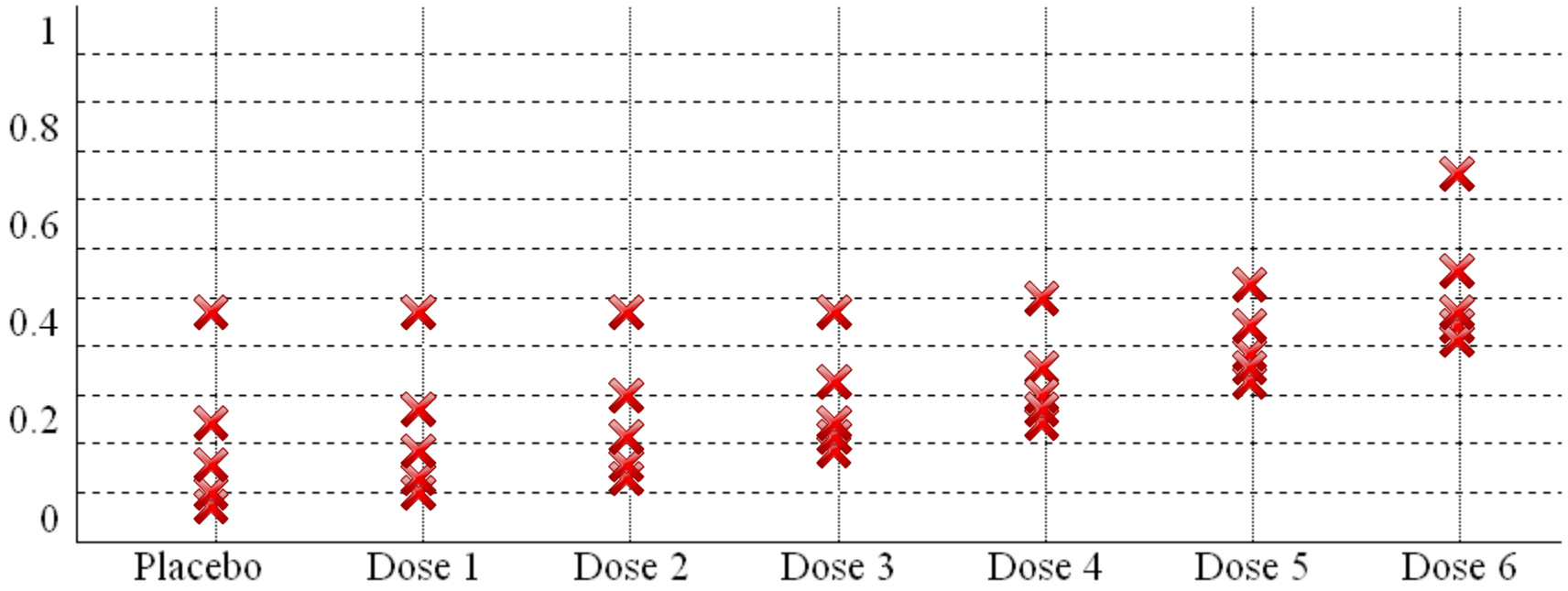


→ *Statements about Posterior Distribution*

→ *Estimates of quantities of interest e.g. ED50*

Imperial College London

Case Study of Huson & Kinnersley



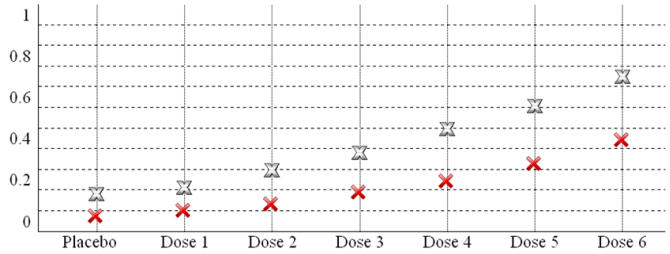
How else could we "combine" these expert beliefs?

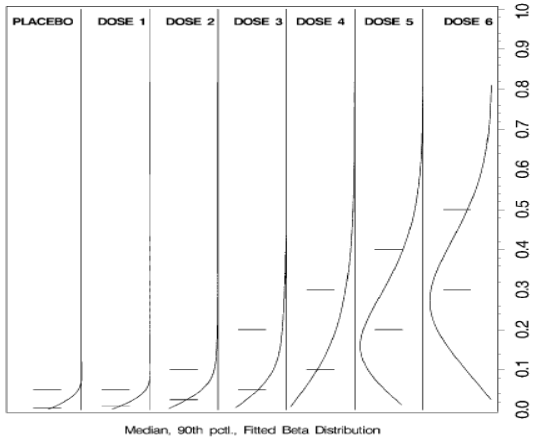
We could bootstrap the expert opinion and form empirical distributions for "a" and "b"

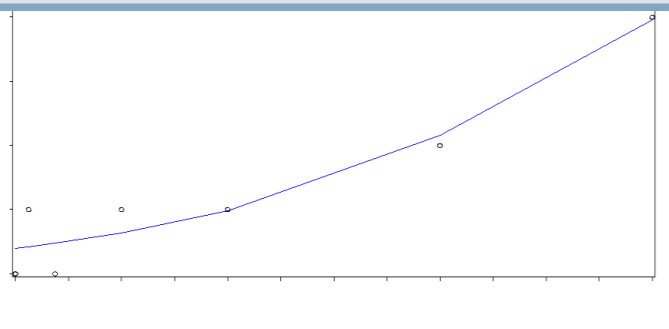
Refinement: 10% of bootstrap samples from "90% sure we can rule out"

Imperial College London

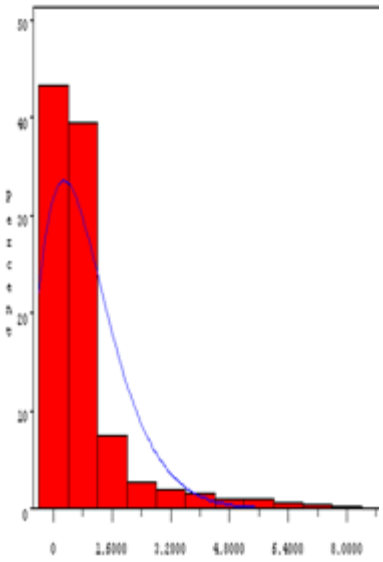
Case Study of Huson & Kinnersley



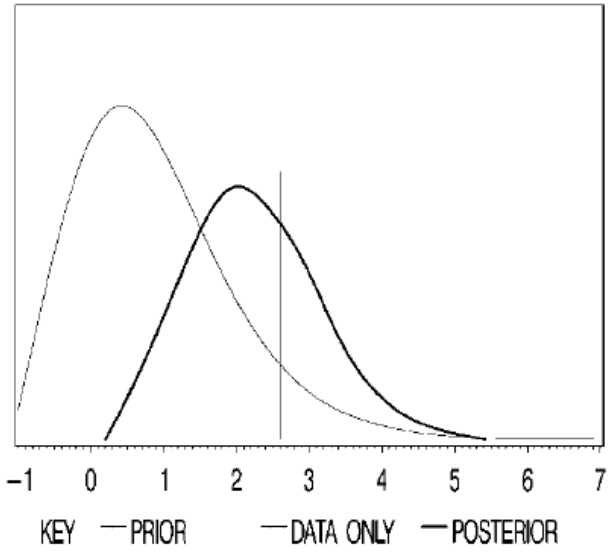




Distn of Slope Parameters & Fitted Beta (to form Prior)



Slope Parameter: Prior, Posterior & Data



- → *Statements about Posterior Distribution*
- → *Estimates of quantities of interest e.g. ED50*

Recent Work in this Area

Huson & Kinnersley (2009)

1. Elicit prior opinion on range of doses
2. Form CDF e.g. beta(α,β)
3. Simulate random values from beta(α,β)
4. Fit logistic curve to simulated values (store 2 parameters: slope, intercept)
5. Repeat simulation & build empirical distributions of slope & intercept
6. Form re-scaled beta distributions for slope & intercept

Also studied: bootstrap sampling instead of steps 2 & 3

Outline

- Bayes- a brief history and pictorial guide
- Example: Statins
- Dose finding in early phase studies
- Risk-benefit decision making

Decision Making- Some background

- A level Maths/ Further Maths – Units D1/2
- Maths BSc module in many universities
- Not routinely part of MSc Medical Statistics training
- Decision-making under uncertainty closely allied with Bayesian statistics for decades, especially in health applications e.g. Raiffa, Schlaiffer, Cornfield, Lindley, Smith AFM, Smith J, Spiegelhalter, Berry, Parmigiani- see Ashby, SiM, 2006 for key references

Evidence Based Medicine

- "EBM is the conscientious explicit, and judicious use of current best evidence in *making decisions* about the care of individual patients" taking into account "individual patients *predicaments, rights and preferences* using *best evidence* from clinically relevant research." Sackett et al, 1996

EBM as Bayesian Decision-Making (Ashby D & Smith AFM, Stats in Medicine, 2000)

- Decision-maker
- Possible actions
- Uncertain consequences
- Sources of evidence
- Utility assessments

Imperial College London

Decision Makers- Who Are they?

- Patients make decisions for themselves, constrained by …
- Prescribing lists of their health care provider who are constrained by …
- NICE who decide on cost-effectiveness, who are constrained by …
- EMEA/ MHRA etc who decide on quality, safety, efficacy and benefit: risk (to individuals and “the public health”), who are constrained by …
- Pharmaceutical companies who decide what to develop and for which licenses to apply

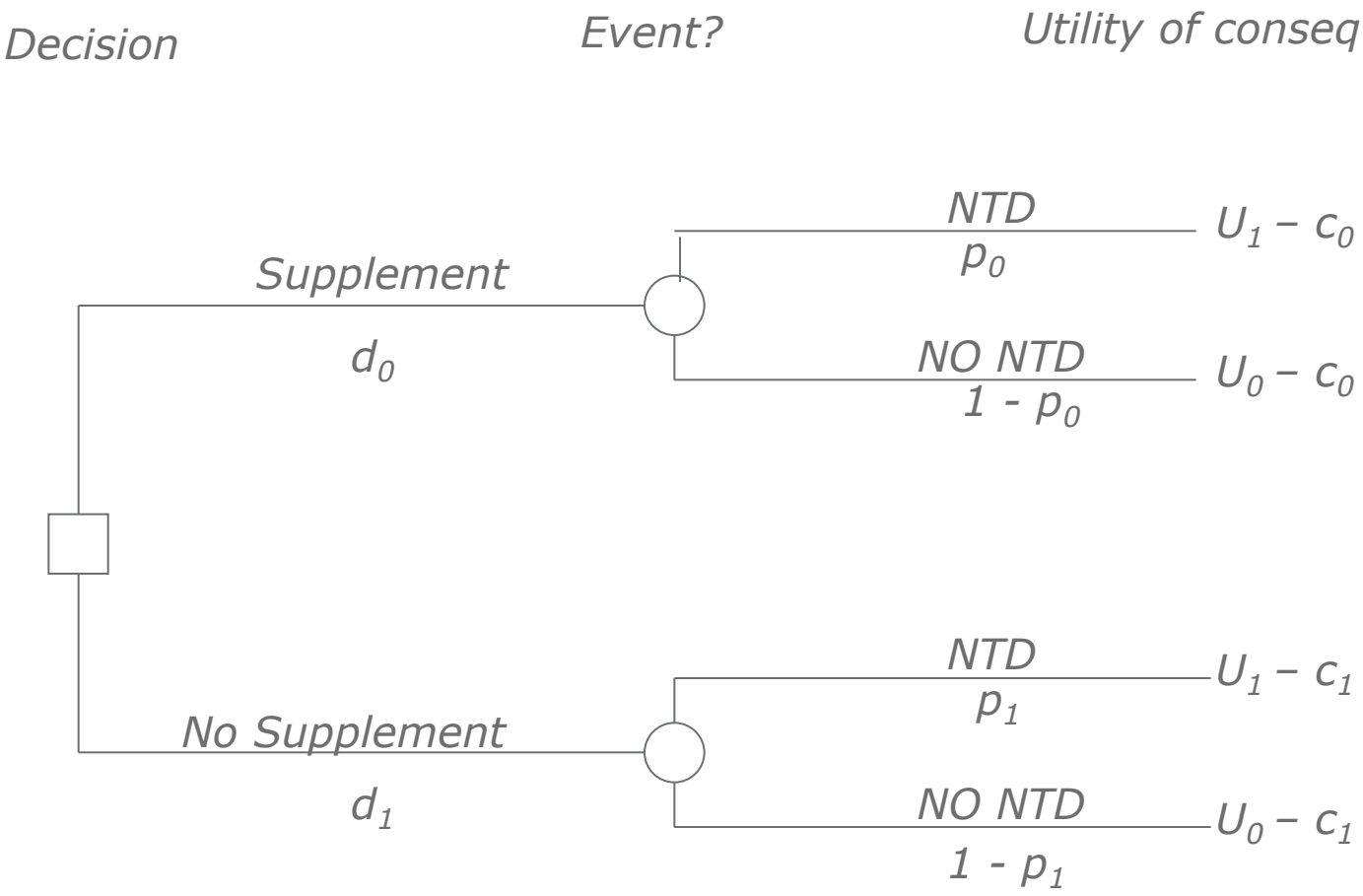
Imperial College London

Couple Wishing to Prevent an NTD

Decision maker	Couple
Possible actions	Take/not take folic acid supplements
Uncertain consequences	Fetus with/without NTD
Sources of evidence	Population statistics Randomised trial in high risk women
Utility assessments	Seriousness of NTD Financial Side-effects

Imperial College London

Model



Imperial College London

Probability Assessment

p_0 *Probability of an NTD with folic acid*
p_1 *Probability of an NTD without folic acid*

4mg in high-risk women:*
p_0 = *10 per 1000 births,* p_1 = *35 per 1000 births*

*low-risk women **:*
p_1 = *3.3 per 1000 births*
p_0 = *1 per 1000 births - assuming 0.4mg folic acid equally effective in low-risk to 4mg folic acid in high-risk*

** Data from MRC Vitamin Study, Lancet 1991*
*** Best estimate from modelling of routine data*

Imperial College London

Probability Assessment

Choose folic acid if $\frac{U_0-U_1}{c_0-c_1} > \frac{1}{(p_0 - p_1)}$

i.e., if $\frac{U_0-U_1}{c_0-c_1} > NNT$

Plugging in previous estimates gives

Previous history of NTD $\frac{U_0-U_1}{c_0-c_1} > 40.3$

No previous history of NTD $\frac{U_0-U_1}{c_0-c_1} > 416.7$

Imperial College London

Public Health Policy on Folic Acid

Decision maker	Minister of Health/CMO
Possible actions	Recommend routine supplementation
Uncertain consequences	Incidence of NTD
Sources of evidence	Population statistics
	Randomised trial in high risk women
Utility assessments	Cost of prescriptions
	Costs of termination/ care
	Desirability of reducing disability

Imperial College London

Herceptin

Benefit: Risk captured with a single parameter

- Pivotal study: randomised, open-label comparing Herceptin and placebo in women with non-metastatic, operable primary invasive breast cancer over-expressing HER2 who had completed … therapy… for primary breast cancer.

- Benefit: Disease-free survival (Placebo vs. Herceptin)
  - proportion with either disease progression or death (due to any cause) 12.9% vs. 7.5%
  - Death (due to any cause) 2.4% vs. 1.8%

- Risk: Cardiotoxicity (Placebo vs. Herceptin)
  - significant asymptomatic (NYHA class I) or mildly symptomatic (NYHA class II) cardiac dysfunction 0.53% vs. 3.04%
  - symptomatic congestive heart failure of NYHA class III or IV or cardiac death 0.06% vs. 0.6%

Imperial College London

Herceptin

Benefit: Risk captured with a single parameter

- MHRA Assessment Report: “If disease-free survival and primary cardiac events **were combined into a single endpoint** it would be dominated by the disease-free survival data with the hazard ratio favouring Herceptin.”

- Benefit: Risk captured with a single parameter assuming equal weight for progression, cardiac event and death from any cause.

- Does further quantification add anything in this type of scenario?

- Could estimate weighting that would need to be given to make the benefit: risk unfavourable, or incidence of cardiac events to make benefit: risk unfavourable given equal weight.

Imperial College London

Treating menopausal symptoms

Decision maker	*Woman*
Possible actions	*HRT or not? For how long?*
Uncertain consequences	*Risk of heart attack/stroke*
	Risk of breast cancer
	Osteoporosis/fractures
	Vasomotor symptoms
	Skin
	Weight change
Sources of evidence	*Epidemiological studies*
	Trials
Utility assessment	*Woman's trade off between long and short term consequences*

Imperial College London

Hormone-replacement therapy: safety update (UK Public Assessment Report, MHRA)

i) 5 years' HRT use in women younger than age 60 years

Type of HRT	*Bsline*	*Absolute risk*	*Attr risk*
Oestrogen-only (no uterus)	*42*	*47 (44–52)*	*5 (2–10)*
Oestrogen-only (w uterus)	*44*	*53 (49–59)*	*9 (5–15)*
Combined HRT	*37*	*51 (48–56)*	*14 (11–19)*

(similar tables for 60-69s, and for 10 years' HRT use)

Imperial College London

Hormone-replacement therapy: safety update (UK Public Assessment Report, MHRA*)

Baseline rate: Obtained by adding the baseline rates for breast cancer, endometrial cancer (in women with a uterus), ovarian cancer, colorectal cancer, venous thromboembolism, CHD, stroke and fracture of femur in non-HRT users.

Absolute risk: Obtained by subtracting the number of cases of colorectal cancer and fracture prevented from the total number of cases of breast cancer, endometrial cancer (in women with a uterus), ovarian cancer, venous thromboembolism, CHD, stroke in HRT users.

Attributable risk: Obtained by subtracting the baseline risk in non-HRT users from the absolute risk in HRT users.

See http://www.mhra.gov.uk/home/groups/pl-p/documents/websiteresources/con2032228.pdf

Imperial College London

Hormone-replacement therapy: safety update (UK Public Assessment Report, MHRA)

"A key drawback of this approach is that the benefits of vasomotor symptom relief—the main indication for HRT—are difficult to quantify and have been not taken into consideration. Because the efficacy of oestrogen-only HRT and combined HRT in relief of vasomotor symptoms is similar, however, the safety profile of these two types of HRT can justifiably be compared."

BUT

- *not very helpful in deciding whether to use HRT or not for its licensed indications*

- *Utilities are implicit- that all other endpoints are equally serious*

cf data-monitoring for WHI (Freedman et al, CCT, 1996; Ashby & Tan, Clinical Trials, 2005)

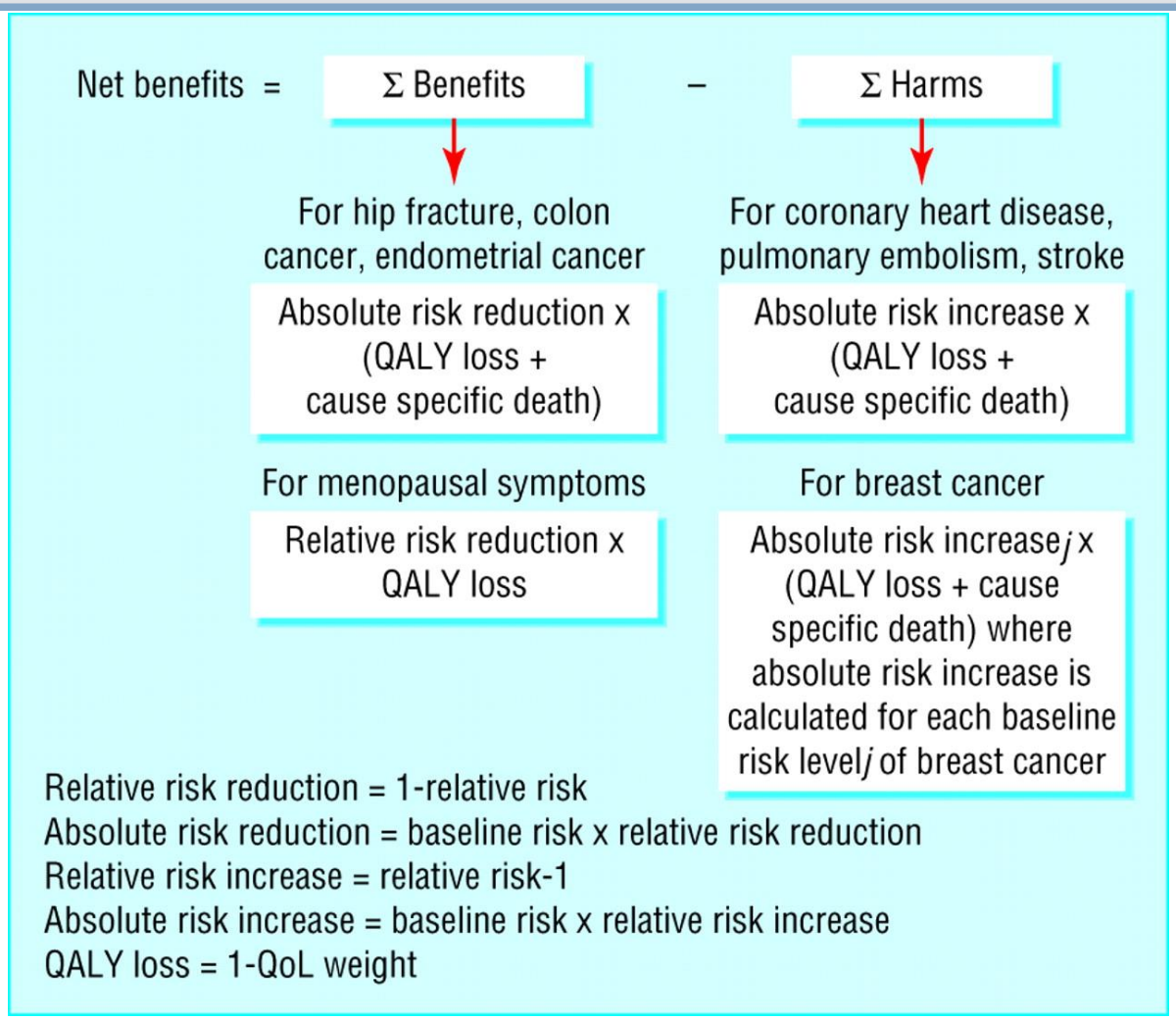
Imperial College London

Benefits and Harms of HRT

(Minelli C et al, BMJ, 2004)

- Objective: to evaluate harms and benefits associated with combined HRT for 5 years for varying baseline breast cancer risk
- Setting: Hypothetical population of white UK women aged 50
- Modelling: Bayesian framework with non-informative priors, fitted via MCMC in WinBUGS based on QALYS and deaths, uses average risks, except for breast cancer
- Data: thoroughly referenced, including HERS I & II, EVTET, WHI

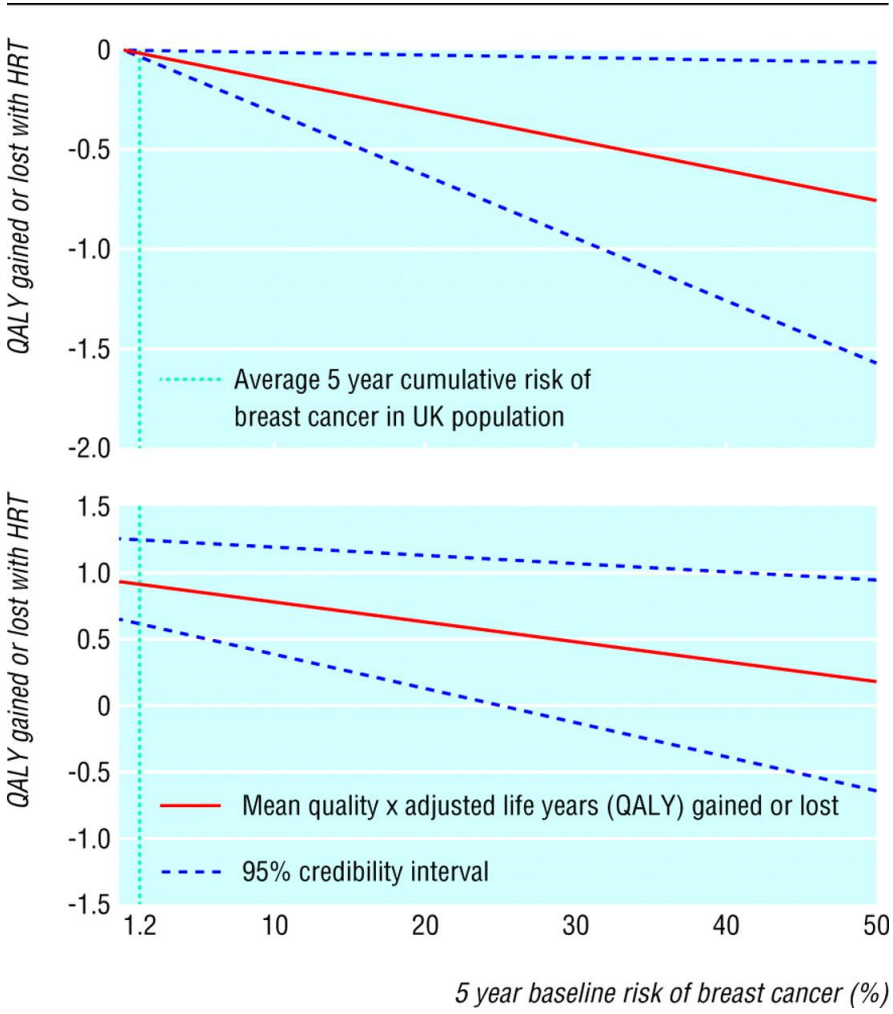
Fig 1 Structure of net benefit decision model



Minelli, C. et al. BMJ 2004;328:371

Imperial College London

Fig 2 Graphical presentation of net-benefit model, with 95% credibility intervals, after exclusion of menopausal symptoms (top) or inclusion of symptoms with QoL weight 0.75 (bottom)

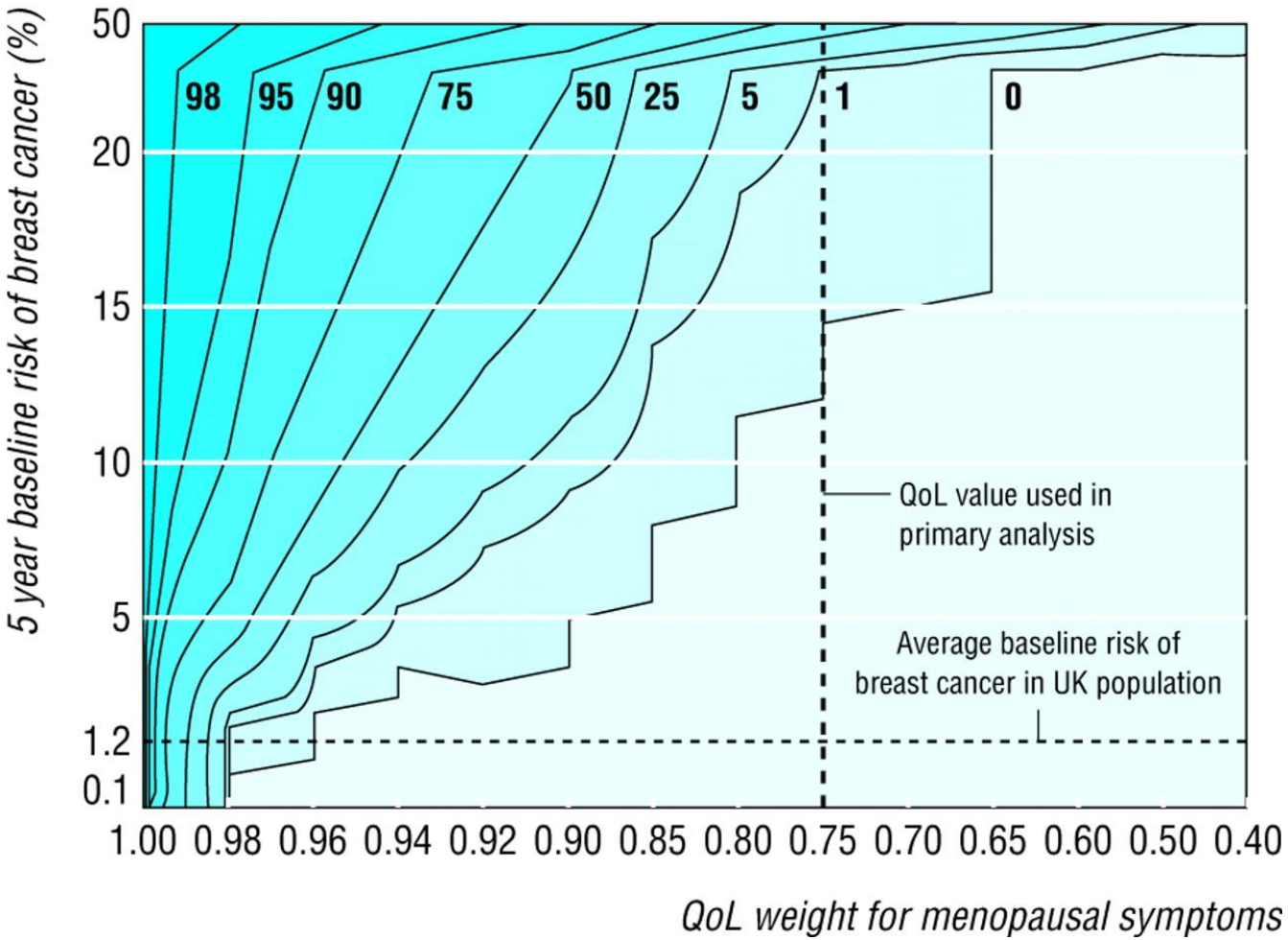


Minelli, C. et al. BMJ 2004;328:371

BMJ

Imperial College London

Fig 3 Probability of net harm (%) associated with HRT use for five years according to utility attributed to menopausal symptoms by individual women and their baseline risks of breast cancer. Isolines define combinations of utility and baseline risk with same probability of net harm



Minelli, C. et al. BMJ 2004;328:371

Imperial College
London

Benefits and Harms of HRT (Minelli C et al, BMJ, 2004)

- Conclusion: “Women with menopausal symptoms on average benefit from HRT,…which concur[s] with the recommendations of the UK MHRA. The results depend on the QoL attributed to symptoms, which in turn vary greatly,….. Thus a decision analysis tailored to individual women would be more appropriate in clinical practice than a population based approach”

Imperial College London

PROTECT-EU

IMI (FP7) Call No 6 “Improving and strengthening the monitoring of the benefit/risk of medicines marketed in the EU” included graphical representation of risk-benefit

PROTECT (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium) led by EMEA and GSK with 29 public and private partners, 2009-2014

Risk-Benefit Decision-Making- lead Imperial

Imperial College London

References

- Ashby D. Bayesian Statistics in Medicine: a 25 year review (2006) *Statistics in Medicine,* **25**, 3580-3631.
- Ashby D**.** and Smith A.F.M. (2000) Evidence based medicine as Bayesian decision-making. *Statistics in Medicine**.*** **19**, 23, 3291-3305.
- Freedman, L., Anderson, G., Kipnis, V., Prentice, R., Wang, C.Y., Rossouw, J., Wittes, J. and DeMets, D. (1996) Approaches to monitoring the results of long-term disease prevention trials: examples from the Women's Health Initiative. *Controlled Clin. Trials*, **17**, 509–525
- Minelli C, Abrams KR, Sutton AJ, Cooper NJ.(2004) Appropriate use of hormone replacement therapy (HRT): A clinical decision analysis. *British Medical Journal* 328: 371-5.
- MRC Vitamin Study Research Group.(1991) Prevention of neural tube defects: results of the MRC vitamin study. 338: 132-7.

References

Accenture Consulting. High performance drug discovery: an operating model for a new era. Accenture (2001)

Huson LW, Kinnersley N. Bayesian fitting of a logistic dose–response curve with numerically derived priors. Pharmaceutical Statistics 2009; 8(4):279-286 DOI: 10.1002/pst.348

ICH E4 (1994). Dose response information to support drug registration. www.emea.eu.int/pdfs/human/ich/037895en.pdf

FDA Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf