An Industry Perspective of the Value of Bayesian Methods

American Course on Drug Development and Regulatory Sciences (ACDRS) Special Workshop: Substantial Evidence in 21st Century Regulatory Science - *Borrowing Strength from Accumulating Data*

David Ohlssen (Statistical Methodology,Novartis Pharmaceutical corporation)

April 21st 2016

NOVARTIS

Acknowledgements

- Björn Bornkamp, Richard Nixon, Beat Neuenschwander, Amy Racine, Satrajit Roychoudhury, Oliver Sander, Heinz Schmidli, Marc Vandemeulebroecke, Andrew Wright

NOVARTIS

Bayesian Thinking in Healthcare Evaluation

Definitions: Spiegelhalter et al. (2004) and Sheiner (1997)

*"The explicit **quantitative** use of **external evidence** in the design, monitoring, analysis, interpretation and reporting of a health-care evaluation"* (Spiegelhalter et al.; 2004)

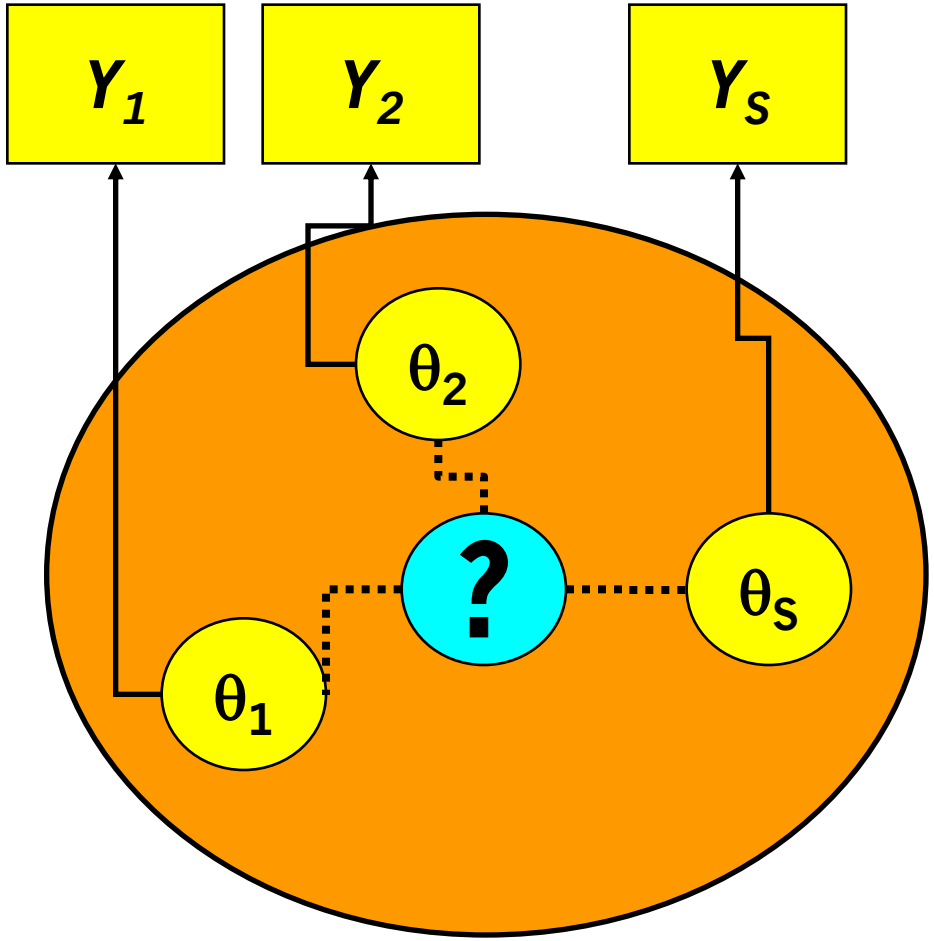
- "…The Bayesian view is well suited to this task because it provides a theoretical basis for learning from experience; that is, for updating prior beliefs in the light of new evidence.
- "I am using the term Bayesian here to describe a point of view, and not a particular statistical method involving use of a prior probability distribution when analysing data. ..."
- "...prior knowledge (i.e., validated scientific theory) is to be incorporated into the analysis of current data, and thereby be updated. Prior knowledge can be introduced, as I stress here, through the assumption of mechanistic scientific models for the data,..."

(Adapted from Learn and Confirm Sheiner;1997)

NOVARTIS

Bayesian Thinking Conceptual Framework

Framework and Notation



$Y_1,..,Y_S$
Data from S sources

$\theta_1,\ldots,\ \theta_S$
Source-specific parameters/effects of interest (*e.g.* a true mean difference)

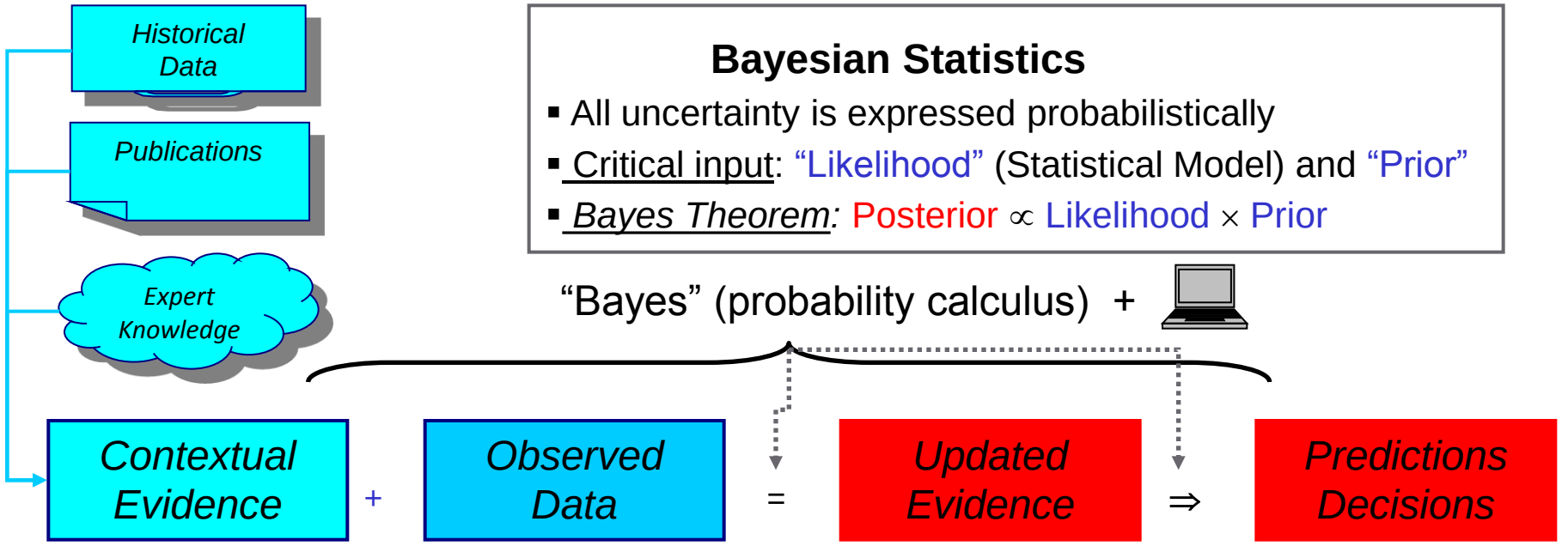
"$(\rightarrow)$ θ *causes* Y"

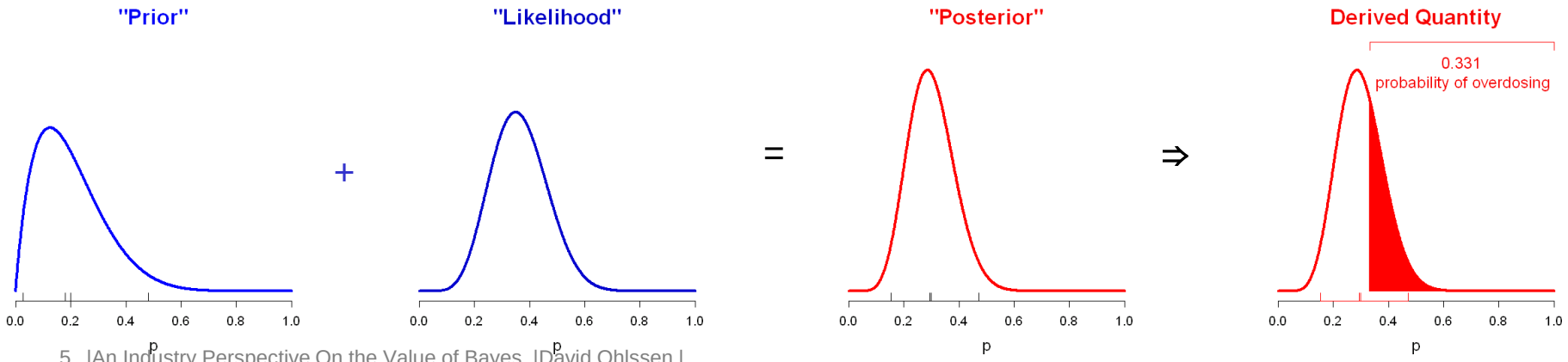
...? ...
Question related to $\theta_1,\ldots,\ \theta_S$ (*e.g.* average effect, effect in a specific study, or effect in a new study)

NOVARTIS

Bayesian Statistics

Summary





Bayesian modeling

Bayesian statistics/methods/models

- Bayesian thinking does not necessarily have to use Bayesian modeling
  - Classical modeling
  - Simulation based techniques
  - *Statistical learning*
  - *Two stage approaches* (set threshold based on first stage)
- However, Bayesian modeling can handle complex settings and incorporates a clear approach to handling and understanding various sources of uncertainty
  - Modeling complexity (non-linear, longitudinal, mechanistic)
  - Full probability models
  - Prediction
  - Decision making

Challenges to using Bayes in Drug Development

- Using Bayes in practice is easier said than done
  - Deciding on what the relevance of different sources of information is subjective and requires scientific expertise
  - Bayesian thinking usually require a much greater level of engagement and resource
  - How to link together relevant evidence and form realistic complex Bayesian models (subjective, requires technical expertise)
- Strong emphasis placed on bias and (strict) type one error control leads to
  - Inference based on one or two pieces of evidence (e.g. confirmatory clinical trials) that are the most rigorous and relevant
  - Being more descriptive and qualitative when assessing other evidence
  - Use of simple methods that focus on population average effects try to avoid models and assumptions

Enabling Bayesian methods with a Structured Framework

- Bayesian statistics often requires a structured framework to be used in practice
- Without a structure it is difficult to convince people you are synthesizing evidence appropriately
  - In Europe, Bayesian methods have been widely used in health technology assessment. However, the backbone of this is a careful systematic review
  - To apply Bayesian methods in benefit risk assessment a structured approach (e.g. multi-criteria decision analysis) is required to identify the key outcomes that should be considered
  - CDRH guidance has greatly helped to provide a structure in trials

Applications

- Bayesian approaches have been used outside of a primary analysis for a confirmatory study
  - Phases I-II and IV (trial design; analysis); Decision making ; Phase III (futility decision rules; missing data sensitivity analysis) ; Integrated safety assessments; Structured benefit risk; Comparative effectiveness and health outcomes ...
- Next we will focus on a few examples
  - Decision making and portfolio assessment
  - Historical data and Meta-analytic predictive priors
  - Design decision making based on posterior probabilities
  - Probability of success

Decision making portfolio assessment

Assess the clinical data in the context of a quantitative TPP

- **Objective:**
  - Support decision making where the aim is to develop a product to meet a medical and market need
  - An approach that is sufficiently flexible to be applied at any stage of drug development across the portfolio
  - Provide quantitative results to stakeholders in a transparent and consistent way
- **Proposed approach:**
  - Define **base case** and **upside quantitative targets** for key efficacy and key safety outcomes in the Target Product Profile (TPP).
  - Identify the **relevant evidence** to assess these **targets**
  - Use probabilities to quantify the current evidence in relation to the TPP targets.
  - Based on a results, align on a common interpretation and a set of recommendations

Overview of the clinical quantitative assessment process

Define TPP and collate the key clinical data

Quantify the TPP

	Base	Upside
Efficacy	0.7	0.5
Safety	0.2	0.1

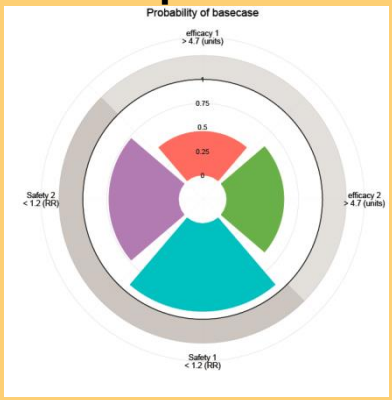
Summarize clinical data

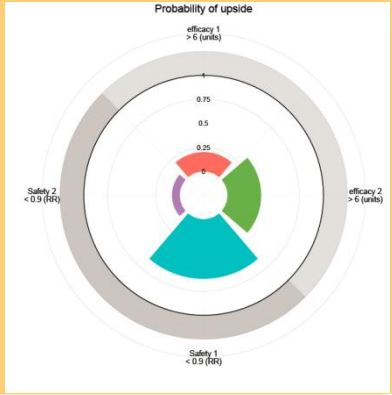
	drug	Standard of care	Diff
Efficacy	1.1	1.5	0.73
Safety	0.3	0.2	0.1

Quantitative assessment of the data

Probabilities quantifying the current evidence

Visual representation





Qualitative assessment of the data

Mirrors the prioritization framework

Generic set of descriptions to assess data

Criteria	Favorable (5)	Neutral (3)	Unfavorable (1)
Efficacy			
Safety			

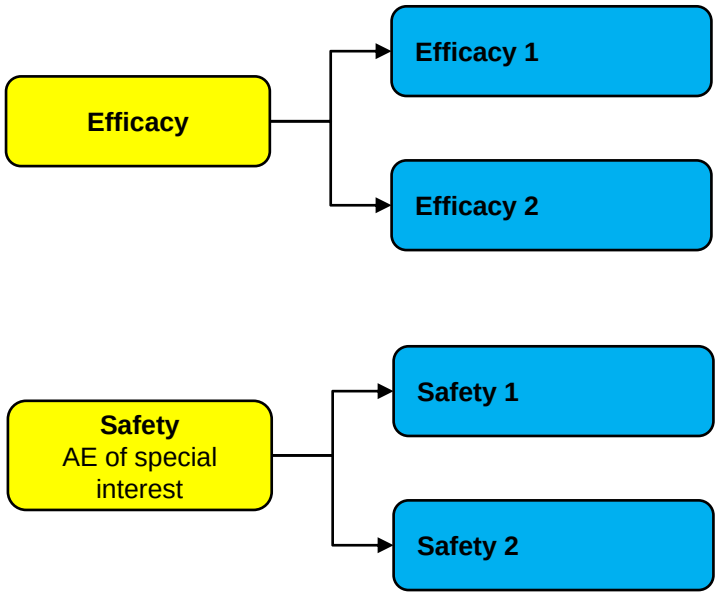
Judgement of the data

Efficacy	(5)	(4)	(3)	(2)	(1)
Safety	(5)	(4)	(3)	(2)	(1)

NOVARTIS

Mock Example: Project ABC data

Create a Value tree to Represent the Quantitative TPP

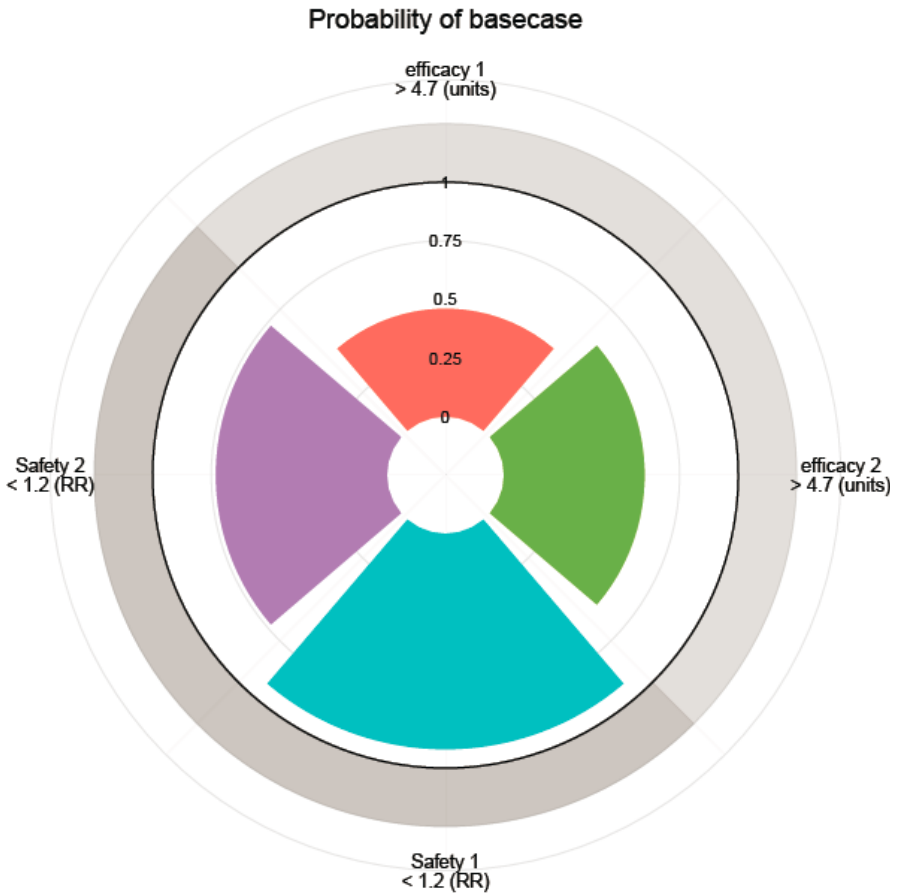


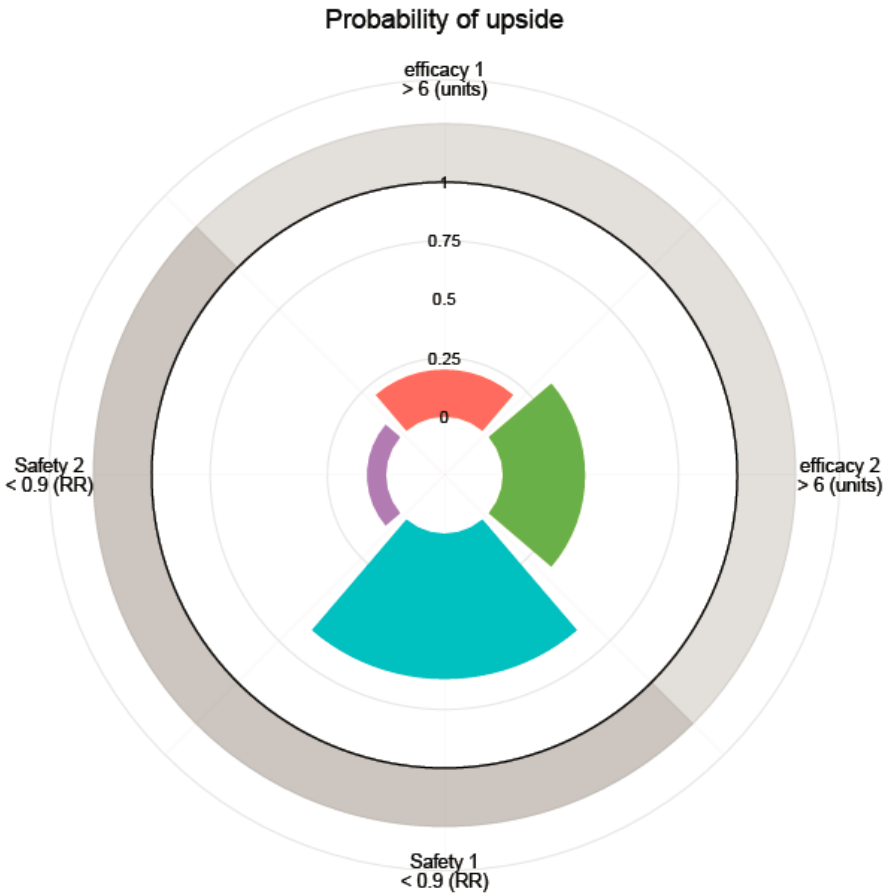
Threshold stat	TPP Base	TPP Upside
4.5 difference (1, 8)	4.7	6
5.2 difference (1.2, 9.2)	4.7	6
0.83 RR (0.4,1.1)	RR 1.2	RR 0.9
1.1 RR (0.85,1.5)	RR 1.2	RR 0.9

NOVARTIS

Mock Example Project ABC123

Summarize key results using a Rose plot





NOVARTIS

Decision Making Portfolio Assessment

Conclusions and The Value Added from the Approach

- Specified quantitative targets leads to an improved TPP
  - This makes sense as all of our projects are judged based on our data (e.g. registration, labeling, comparative effectiveness)
  - The team discussion to develop these targets leads to a stronger link between clinical development and commercial objectives.
- Puts evidence for multiple factors on the same scale
  - Enables the link of evidence from very different end-points to the targets
  - Provides a consistent picture of all relevant data of a given project at a given time point
  - Leads disciplined approach to decision making based on the evaluation of the current evidence related to clinical development and commercial objectives
  - Uncertainty assessed by using two targets (base and upside cases)

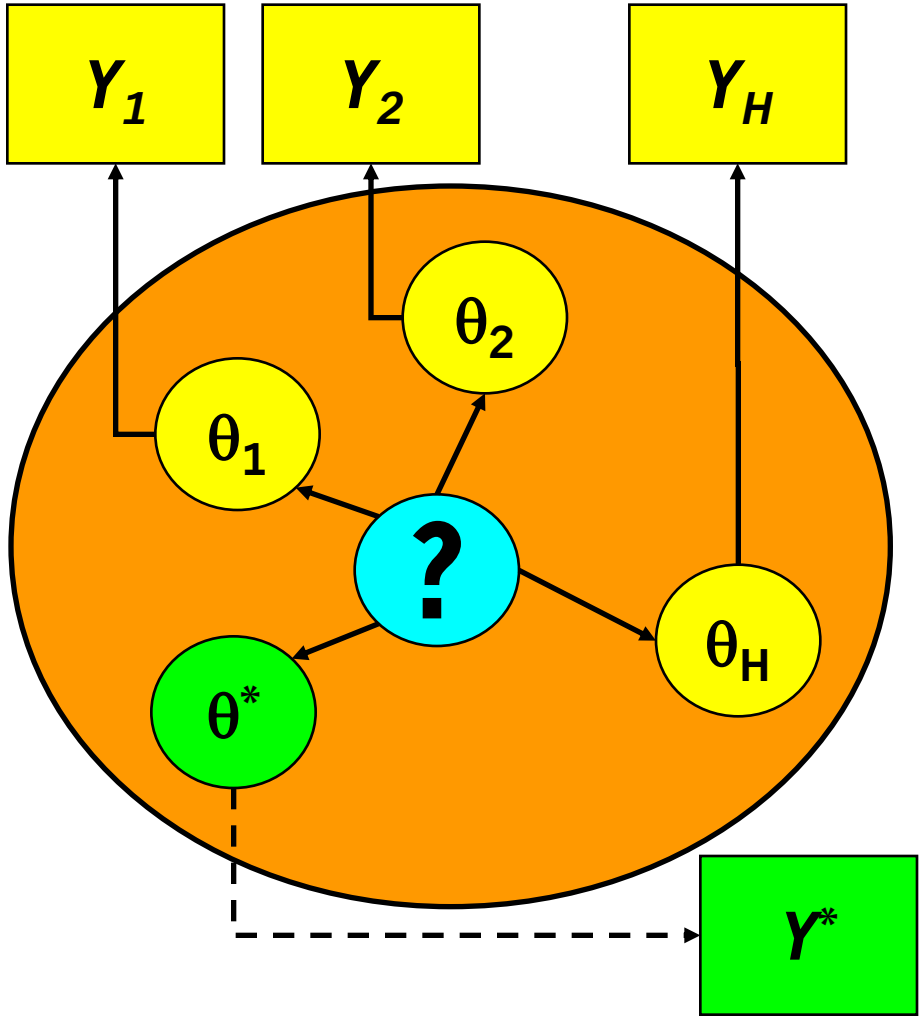
Using Historical Control Data

Objective and Problem Statement

- Design a study with a control arm / treatment arm(s)
- Use historical control data in design and analysis
- Ideally: $\rightarrow$ smaller trial comparable to a standard trial
- Used in some of Novartis phase I and II trials
- Design options
  - Standard Design: "*n* vs. *n*"
  - New Design: "*n**+*(n-n*)* vs. *n*" with *n** = "prior sample size"
- How can the historical information be quantified?
- How much is it worth?

The Meta-Analytic-Predictive Approach

Framework and Notation



$Y_1,..,Y_H$
Historical control data from
H trials

$\theta_1,\ldots,\ \theta_H$
Control "effects" (unknown)

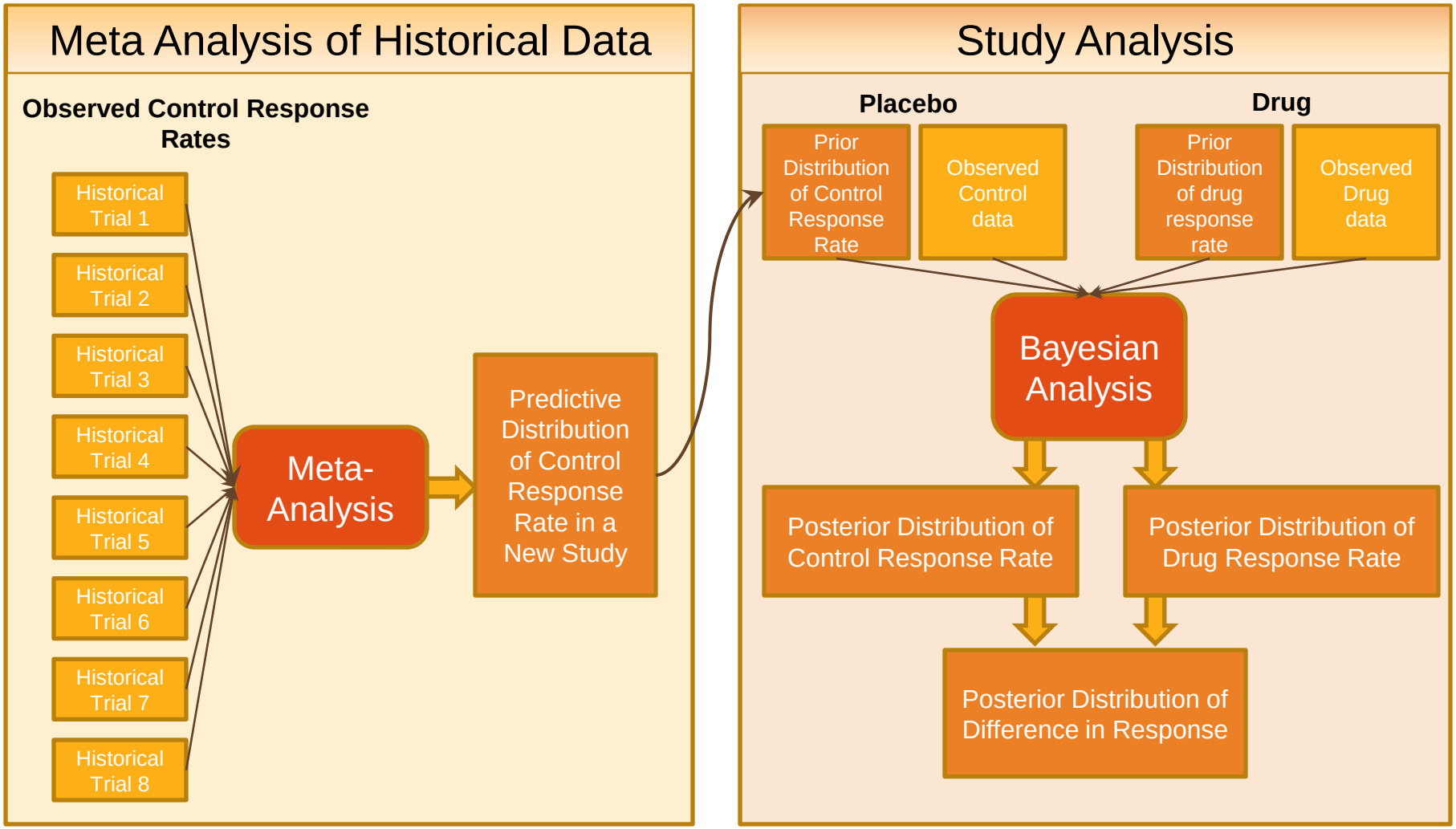
?
'Relationship/Similarity'
(unknown)
no relation ←…→ same effects

θ^*
Effect in new trial (unknown)
Design objective: $[\ \theta^* \mid Y_1,\ldots,Y_H\ ]$

Y^*
Data in new study
(yet to be observed)

NOVARTIS

Bayesian setup using historical control data



NOVARTIS

Example Ankylosing Spondylitis Study

Application in of using historical control data in a Proof of Concept Study

- *Disease*
  Ankylosing spondylitis
- *Experimental treatment*
  Monoclonal antibody
- *Endpoint*
  Binary: response at week 6
- *Traditional clinical trial design*
  - Experimental (n=24) vs. Placebo (n=24)
  - Fisher's exact test

However: 8 similar historical placebo-controlled clinical trials with different experimental treatments available

Could this historical placebo information be used?

Historical Controls

Motivating example: Trial design and analysis with historical controls

Historical placebo information

- Bayesian primary analysis
- *Prior Placebo* Derived from 8 historical trials (N=533), using a Meta-Analytic-Predictive (MAP) approach

  Beta(11,32) worth 43=11+32 patients
- *Prior Experimental* Weakly informative

  Beta(0.5,1) worth 1.5=0.5+1 patients
- Design:

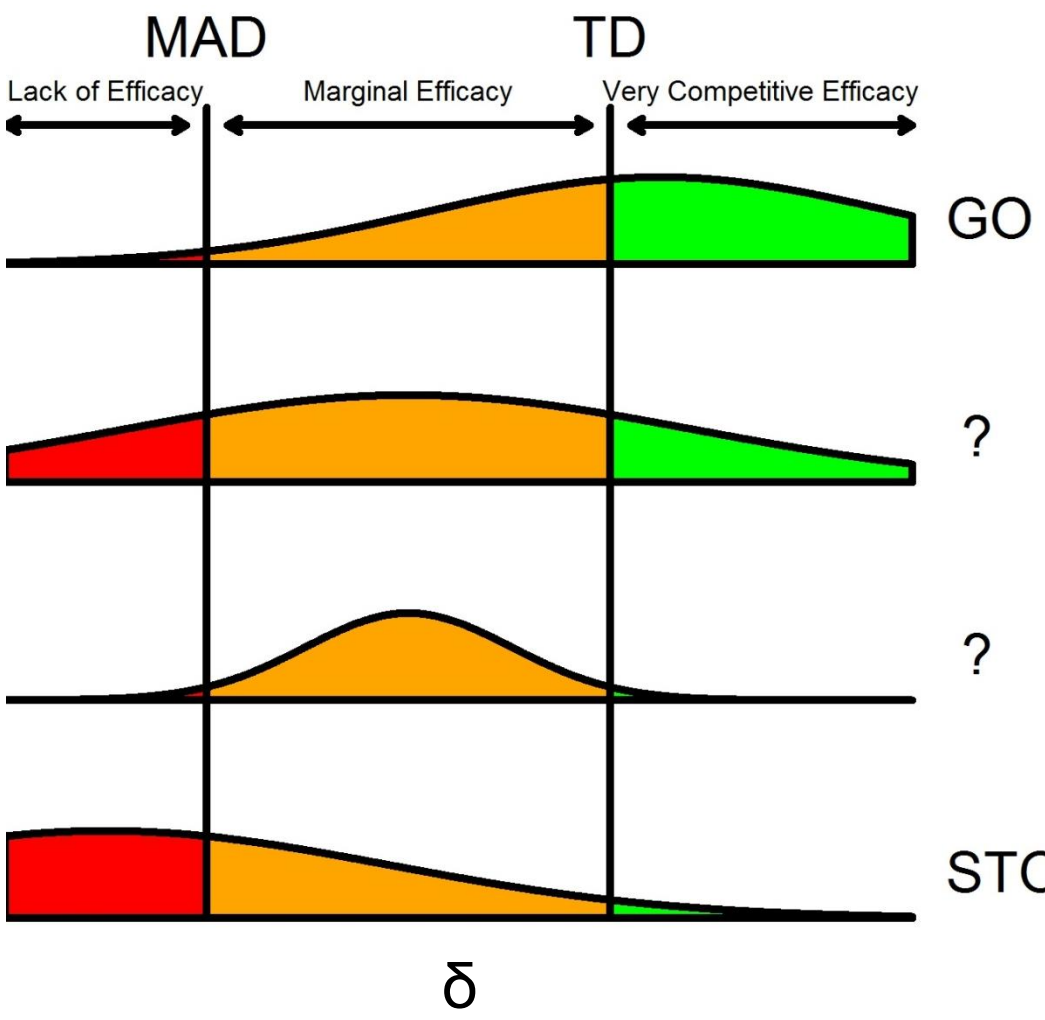
  Secukinumab (n=24) vs. Placebo (n=6)
- Results:

  14/24 Secukinumab vs. 1/6 Placebo, $p(\delta > 0 \mid \text{Data}) > 99.8\%$

  Baeten et al. (2013) Lancet 382(9906):1705-1713

Decision rules based on Posterior Probability

Double criterion - minimal acceptable difference target difference



Treatment vs. Control

p(δ > MAD | data) > 97.5%

p(δ > TD | data) > 50%

indeterminate:

neither STOP nor GO

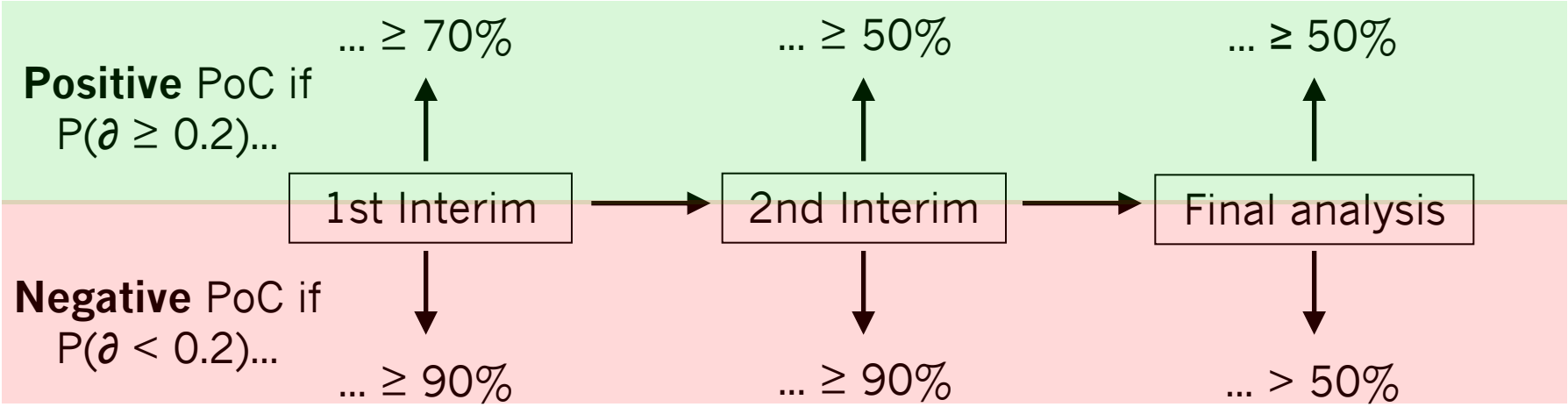
p(δ < MAD | data) > 50%

p(δ < TD | data) > 80%

NOVARTIS

Utilization in a Quick kill Quick win PoC Design

Assessing the design using Frequentist Operating Characteristics



With N=60, 2:1 Active:Placebo, IA's after 20 and 40 patients

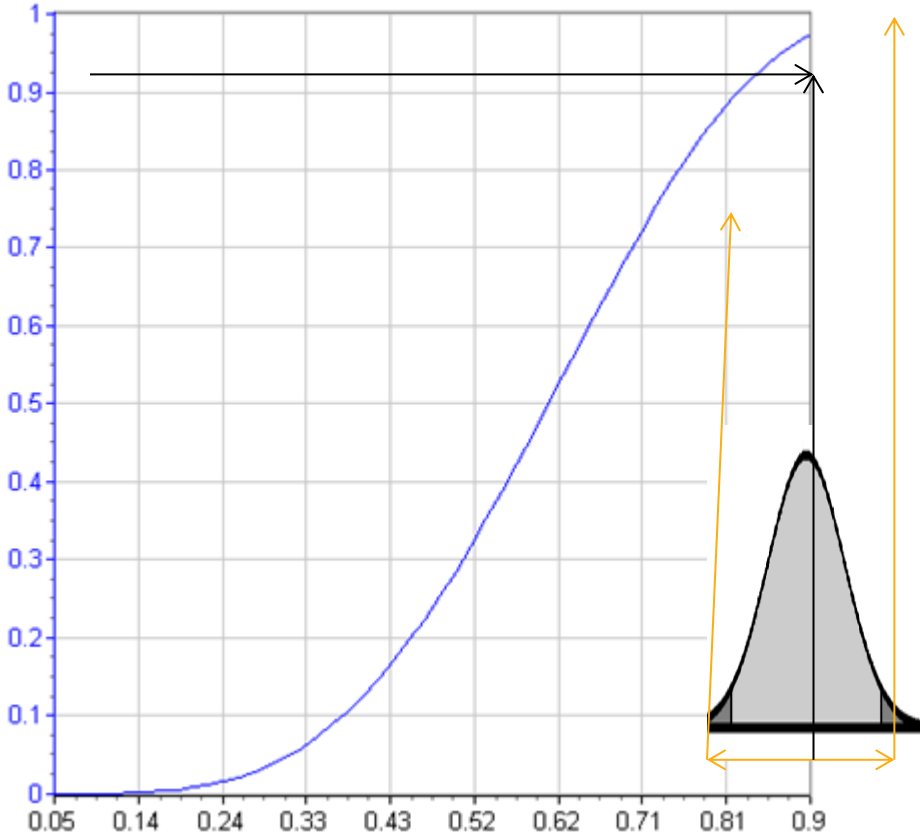
	First interim		Second interim		Final		Overall power
Scenario	Stop for efficacy	Stop for futility	Stop for efficacy	Stop for futility	Claim efficacy	Fail	
∂ = 0	1.6%	49.0%	1.4%	26.0%	0.2%	21.9%	3.2%
∂ = 0.2	33.9%	5.1%	27.7%	3.0%	8.8%	21.6%	70.4%
∂ = 0.5	96.0%	0.0%	4.0%	0.0%	0.0%	0.0%	100.0%

With $p_{Placebo}$ = 0.15, 10000 runs

NOVARTIS

Introduction to Probability of success

- Predictive distributions can be used to calculate the probability a future study will be successful based on the results of a previous study
- Standard power calculations will assume fixed treatment difference. Based on this, a phase III sample size would be chosen to achieve, say, 90% power
- Probability of success (PoS) calculates the probability of phase III success, given a phase III design, accounting for uncertainty surrounding the treatment effect assumption
- Typically, PoS will be lower than the specified power



NOVARTIS

PoS End of phase II sensitivity analysis

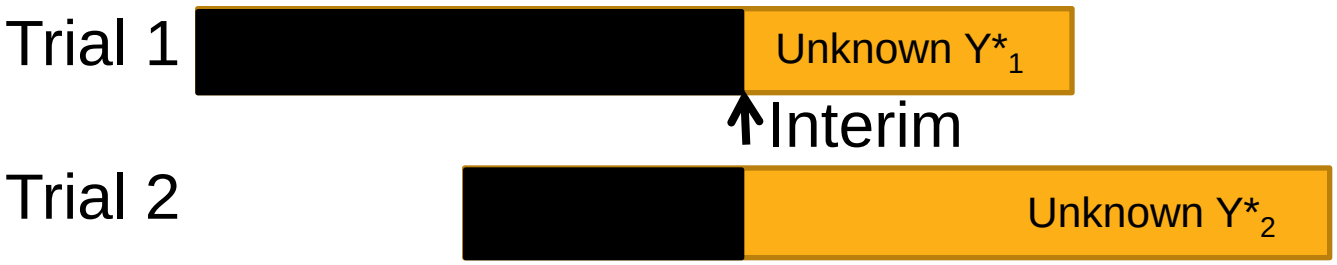
Accounting for multiple Ph III outcomes dosing strategies and NI margins

	Both Co-primary	Accounting for tolerability
BASE CASE	**64%**	**52%**
Select alternative dose with lower efficacy	**-14%**	**-3%**
Alternative dose equal efficacy	**+0%**	**+11%**
Tougher NI margin outcome 1	**-8%**	**-4%**
Tougher NI margin Outcome 2	**-3%**	**-3%**

NOVARTIS

Two ongoing phase 3 trials, one delayed

Two “identical“ phase 3 trials (almost same centers)



Success

p-value <0.05 in one, <0.1 in other

Predictive distribution: continuous primary endpoint Y

- All patients recruited: baseline covariate X known
- Bivariate normal assumption (Y, X) (by treatment)

NOVARTIS

Discussion

- Within some companies Bayesian methods are reasonably widely used for internal decision making.
- Frameworks, such as CDRH guidance and UK NICE approach to HTA assessment, have helped move Bayesian methods into regulatory decision making
- Frameworks under development for extrapolation (e.g. adults to pediatrics), structured benefit risk and safety meta-analysis might lead to wider use of Bayesian methods
- Bayesian thinking is more important than Bayesian statistics
- A pragmatic approach, with emphasis on addressing the problem, using elements of both Bayesian and frequentist methods, is recommended

References

Weblinks

- Cochrane library www.cochranelibrary.com
- ISPOR www.ispor.org/workpaper/practices_index.asp
- NICE www.nice.org.uk/guidance
- CADTH www.cadth.ca/indirect-evidence-indirect-treatment-comparisons-meta-analysis
- EMA www.ema.europa.eu/ema
- FDA www.fda.gov

References

Books

- Spiegelhalter DJ, Abrams KR, Myles JP (2004) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation.* Wiley.
- Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades AE (2012) *Evidence Synthesis for Decision Making in Healthcare.* Wiley.
- Egger M, Davey-Smith G, Altman D – editors (2001) *Systematic Reviews in Health Care: Meta-Analysis in Context.* Wiley.
- Higgins JPT, Green S – editors (2011). *Cochrane Handbook for Systematic Reviews of Interventions.* The Cochrane Collaboration.
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) *Bayesian Data Analysis.* CRC Press.
- Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D (2012) *The BUGS Book: A Practical Introduction to Bayesian Analysis.* CRC Press.

References

Evidence Synthesis

- Higgins JP, Whitehead A (1996) Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine* 15, 2733-2749.
- Sutton AJ, Abrams KR (2001) Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research* 10, 277-303.
- Lumley T (2002) Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* 21, 2313-24.
- Lu G, Ades AE (2004) Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 23,3105-3124.
- Ohlssen DI, Sharples LD, Spiegelhalter DJ (2007) Flexible random-effects models using Bayesian semi-parametric models: applications to institutional comparisons. *Statistics in Medicine* 26, 2088-112.
- Salanti G, Higgins JP, Ades AE, Ioannidis JP (2008) Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* 17, 279-301.

References

Historical Controls

- Pocock SJ (1976) The combination of randomized and historical controls in clinical trials. *Journal of Chronic Diseases* 29, 175-188.
- Ibrahim JG, Chen MH (2000) Power prior distributions for regression models. *Statistical Science* 15, 46-60.
- Neuenschwander B, Capkun-Niggli G, Branson M, Spiegelhalter DJ (2010) Summarizing historical information on controls in clinical trials. *Clinical Trials* 7, 5-18.
- Hobbs BP, Carlin BP, Mandrekar SJ, Sargent DJ (2011) Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. *Biometrics* 67, 1047-1056.
- Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 70(4):1023-1032.

References

Probability of Success

- Spiegelhalter DJ, Freedman LS, Blackburn PR (1986) Monitoring clinical trials: conditional or predictive power? *Control Clinical Trials* 7, 8-17.
- O'Hagan A, Stevens JW, Campbell MJ (2005) Assurance in clinical trial design. *Pharmaceutical Statistics* 4, 187-201.

Additional reference

- Sheiner, L.B., (1997). Learning versus confirming in clinical drug development. Clin. Pharmacol. Ther. 61, 275–291