Bayesian Methods for Clinical Studies

Emmanuel Lesaffre
Interuniversity Institute for Biostatistics and statistical Bioinformatics
KU Leuven & University Hasselt
emmanuel.lesaffre@kuleuven.be

EJP RD_WP 20 Intermediate course
March 22, 2024

A crash course — Webinar

Are you a Bayesian?

But ...

0.1 We think Bayesian – in our natural life

Bayesian approach mimics our natural life where learning is done by combining past and present experience

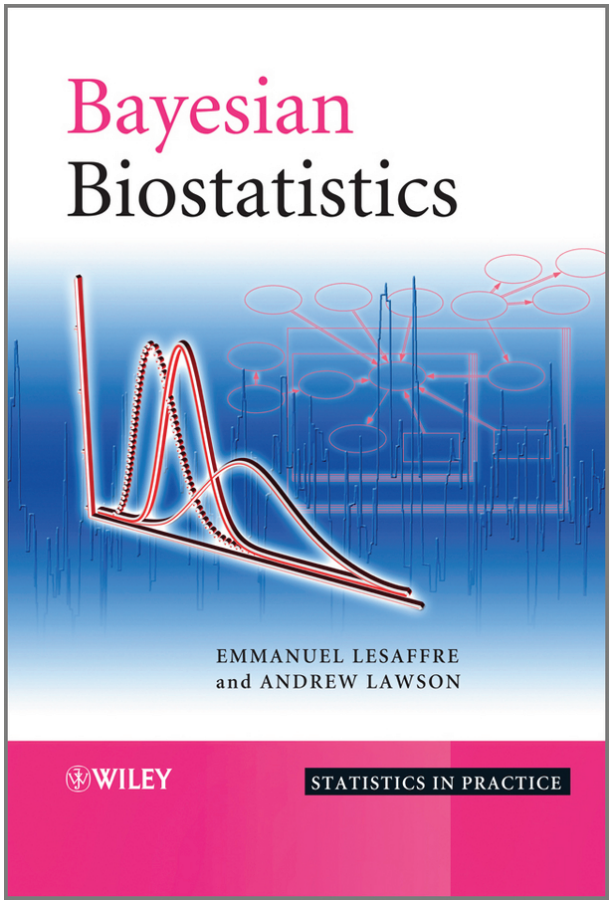
You don't consult the car's manual, each time you drive your car!

0.2 We act Bayesian – in our research

A significant result on a small trial

- Small sized study with an unexpectedly positive result about a new medication to treat patients with oral cancer
- First reaction (certainly of the drug company) = "**great!**"
- Past: none of the medications had such a large effect and new medication is not much different from the standard treatment
- Second reaction (if one is honest) = **be cautious**
- Then, **You are a Bayesian (statistician)**

Part of the material is obtained from



Session 1: A crash course in Bayesian methodology

Session 1
A crash course in Bayesian methodology

- What is the Bayesian approach?
- What are the differences with the classical statistical approach?
- Some simple illustrations of the Bayesian approach

Central idea of Bayesian approach

Combine data (likelihood) with Your prior knowledge (prior probability) to update information on the parameter to result in a revised probability associated with the parameter (posterior probability)

1.1 Bayes theorem – Basic version

From $\Pr(A, B) = \Pr(B \mid A)\Pr(A) = \Pr(A \mid B)\Pr(B)$

- **Basic version of Bayes Theorem**:

$$\Pr(B \mid A) = \frac{\Pr(A \mid B) \times \Pr(B)}{Pr(A)}$$
$$= \frac{\Pr(A \mid B) \times \Pr(B)}{\Pr(A \mid B) \times \Pr(B) + \Pr(A \mid B^C) \times \Pr(B^C)}$$

- Example: predictive value $\Leftrightarrow$ sensitivity, specificity, prevalence of disease
  - $A =$ "positive diagnostic test"
  - $B =$ "diseased"

Sensitivity, specificity, prevalence and predictive values

- $A =$ "positive diagnostic test", $B =$ "diseased"
- Characteristics of diagnostic test:
  - ▷ Sensitivity $(S_e) = \Pr(A \mid B)$
  - ▷ Specificity $(S_p) = \Pr(A^C \mid B^C)$
  - ▷ Positive predictive value (pred+) $= \Pr(B \mid A)$
  - ▷ Negative predictive value (pred-) $= \Pr(B^C \mid A^C)$
  - ▷ Prevalence (prev) $= \Pr(B)$
- $pred+$ calculated from S_e, S_p and $prev$ using Bayes theorem

▷ Well-known example of **Bayes Theorem**: Predictive values of a diagnostic:

$$pred+ = \frac{S_e \cdot prev}{S_e \cdot prev + (1 - S_p) \cdot (1 - prev)}$$

▷ Thus, given:

- sensitivity: probability of positive diagnostic test if diseased
- specificity: probability of negative diagnostic test if healthy
- prevalence of disease: proportion of subjects in population that are diseases

▷ Bayes Theorem allows to compute:

- positive predictive value: probability of diseased if positive diagnostic test
- negative predictive value: probability of healthy if negative diagnostic test

for any value of the prevalence

Predictive value of the OralCDx technique

- ▷ OralCDx technique: computerized analysis of brush biopsies, to detect dysplasia or carcinoma in patients with oral mucosal lesions (disease)
- ▷ Source: Chapter 12 of Alonso and Gianninni in Lesaffre, Feine, Leroux and Declerck (2009)
- ▷ D=1/0 (disease present/absent), T=1/0 (positive/negative OralCDx results)

	D=0	D=1	
$T = 0$	68	10	78
$T = 1$	2	16	18
	70	26	96

- ▷ For this table:
  - $S_e = 16/26 = 0.62$, $S_p = 68/70 = 0.97$, $prev = 26/96 = 0.27$, apparent $prev = 18/96 = 0.19$
  - pred+ $= 16/18 = 0.89$ & pred- $= 0.87$
- ▷ If prevalence changes, use Bayes Theorem, e.g. for prev $= 0.03, 0.10$, or ...

▷ Prior (prevalence): belief that patient has dysplasia or carcinoma in population

▷ Data (likelihood): OralCDx result (positive or negative)

▷ Posterior (pred+ or pred-): combination of prevalence with test result

$\Rightarrow$ Posterior $\propto$ prior $\times$ likelihood

1.2 Bayes theorem – General expression

- ▷ We now derive the general expression of Bayes Theorem
- ▷ It is the basis for the increasingly popular Bayesian approach
- ▷ Suppose we compare two treatments expressed by an unknown parameter θ:
  - ◦ Binary outcome: odds ratio, relative risk, absolute risk reduction
  - ◦ Continuous outcome: difference in means
  - ◦ Count: rate ratio
- ▷ For a Bayesian analysis we need:
  - ◦ to combine prior information on θ
  - ◦ with information obtained on θ from current data
  - ◦ to obtain a posterior idea on θ

Advantages of Bayesian approach over classical statistical approach

- ▷ External information can be formally incorporated
- ▷ Output of a statistical exercise is formulated in a much more elegant manner
- ▷ No need to imagine what could have happened as in the calculation of the P-value
- ▷ Bayesian software allows to fit (much more) complex models

This is what I wish to show!

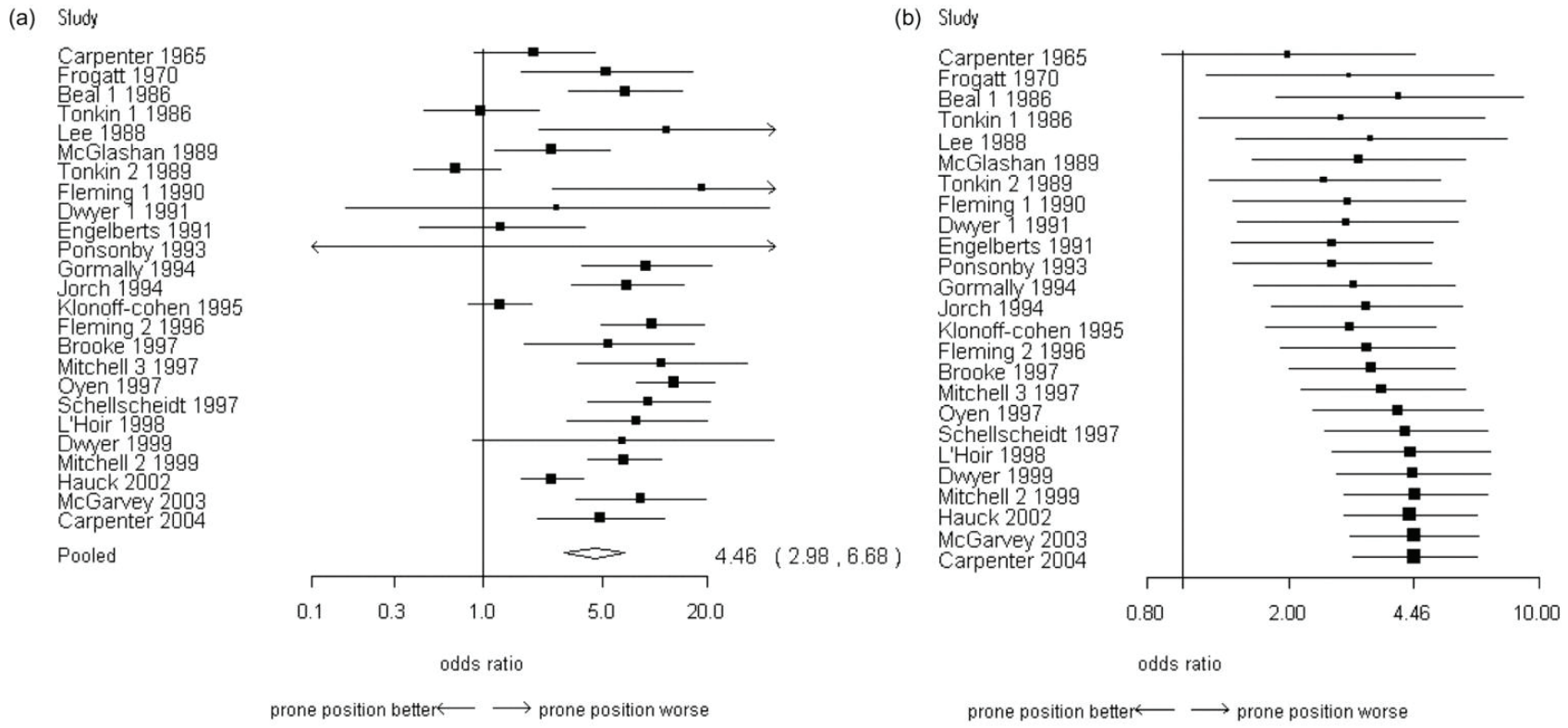
To get an idea, the following story on SIDS may help!

Infant sleeping position and SIDS

▷ **Question**: Should baby sleep in **front** or **back** position to avoid S(udden) I(nfant) D(eath) S(yndrome)?

?

▷ Meta-analysis by Gilbert et al. (2005)



- Conclusion from **cumulative meta-analysis**:
  - ▷ Advice to put infants to sleep on the front for nearly a half century was contrary to evidence available from 1970 that this was likely to be harmful.
  - ▷ A systematic review of preventable risk factors for SIDS from 1970 would have led to earlier recognition of the risks of sleeping on the front and might have prevented over 10 000 infant deaths in the UK and at least 50 000 in Europe, the USA, and Australasia.

- **Now suppose the following**:
  - ▷ I am **Tonkin** and **I am preparing my paper on SIDS** in **1985**
  - ▷ I do a classical analysis of the obtained data, which are:

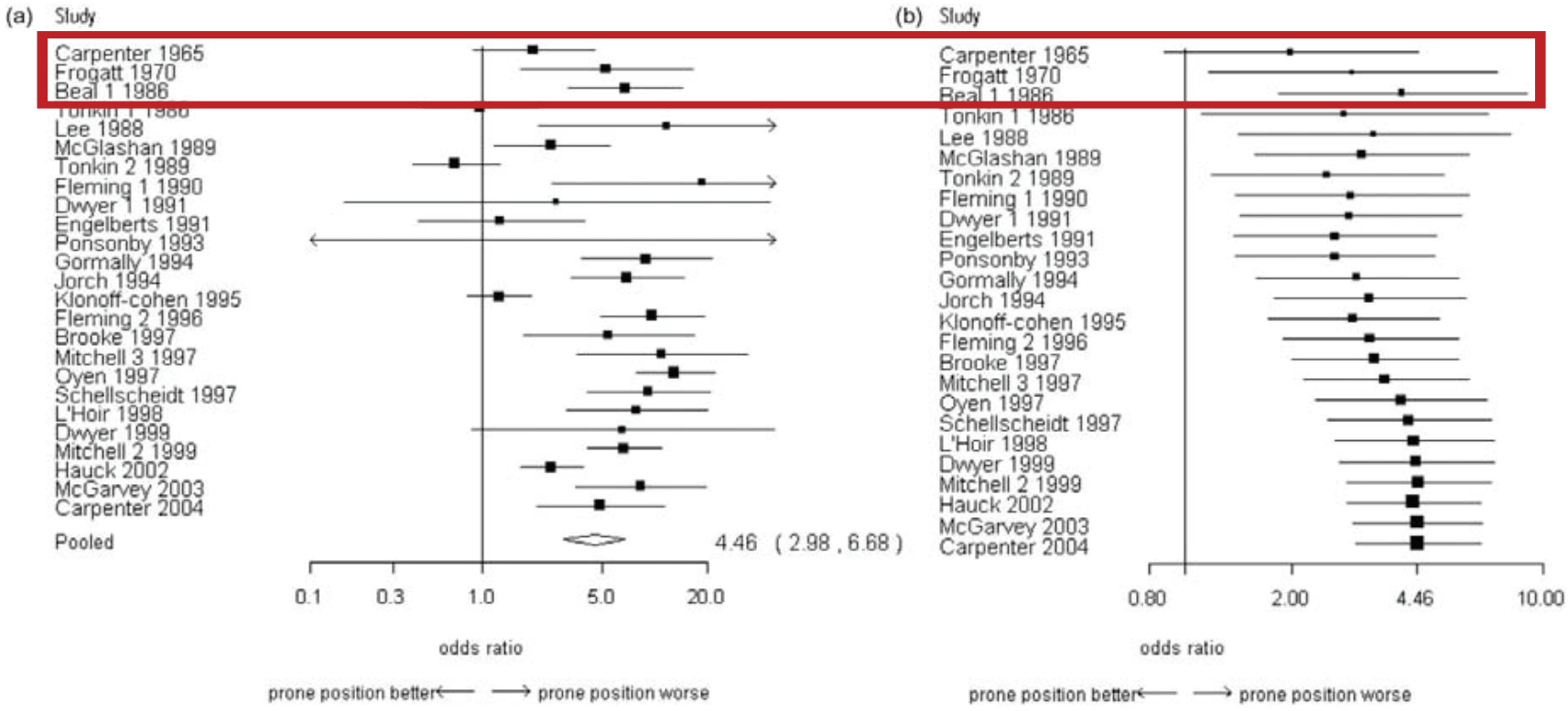
	SIDS=No	SIDS=Yes	
Front	575	51	626
Back	79	8	87
	654	59	713

  - ▷ Using a standard statistical program, I obtain for OR $(= \theta)$:

    $\widehat{OR} = 0.88$ with 95% CI $=$ [0.40, 1.90] and P $=$ 0.74

- **Suppose later on**:
  - ▷ I picked up the result of the cumulative meta-analysis after the first 3 studies
  - ▷ Roughly I observe from the figure $OR_0 = 4.20$ with 95% CI $=$ [1.9, 9.3]
  - ▷ **What can I do with this prior information?** ... Some possibilities:
    1. Ignore the prior information
    2. Combine the prior information with the current data
    3. Combine discounted prior information with the current data
  - ▷ I have already done case 1
  - ▷ But how to do cases 2 and 3?
  - ▷ We (can) use **Bayesian approach**

▷ Meta-analysis by Gilbert et al. (2005)



Information from first three studies highlighted

▷ Three Bayesian analyses provide estimates of OR for SIDS in front position

	PRIOR INFORMATION on ODDS RATIO		
	Nothing	**Cum MA**	**Reduced**
Prior	1 [0, 10^6]	4.20 [1.9, 9.3]	4.20 [1.58, 11.19]
Data	0.88 [0.40, 1.90]	0.88 [0.40, 1.90]	0.88 [0.40, 1.90]
Posterior	0.88 [0.38, 1.81]	2.09 [1.11, 3.63]	1.74 [0.86, 3.39]

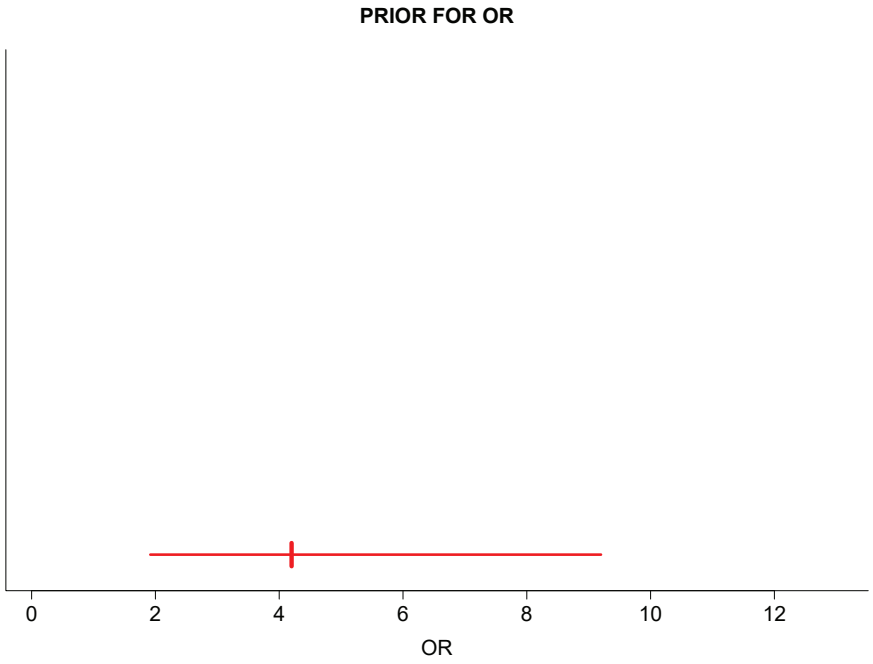
▷ Findings of Bayesian analyses:

- When nothing is assumed, then result of (my) Tonkin study is obtained
- When results from previous 3 studies are taken into account, risk in front position for SIDS ≈ as in a standard cumulative MA
- When prior information on risk is reduced, then OR is pulled towards 1 but is still larger than 1

How to specify prior information on θ?

▷ Information from Carpenter (1965) & Frogatt (1970) & Beal (1986):
$\widehat{OR} = 4.20$ with 95% CI $= [1.90, 9.30]$

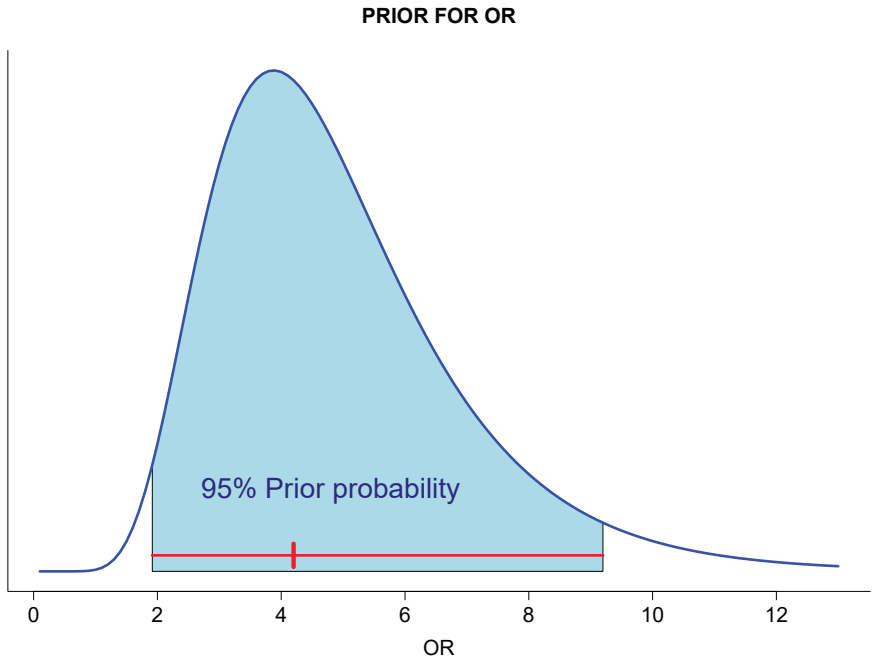
Information from first 3 studies on SIDS



How to specify prior information on θ?

▷ Information from Carpenter (1965) & Frogatt (1970) & Beal (1986):
$\widehat{OR} = 4.20$ with 95% CI $= [1.90, 9.30]$

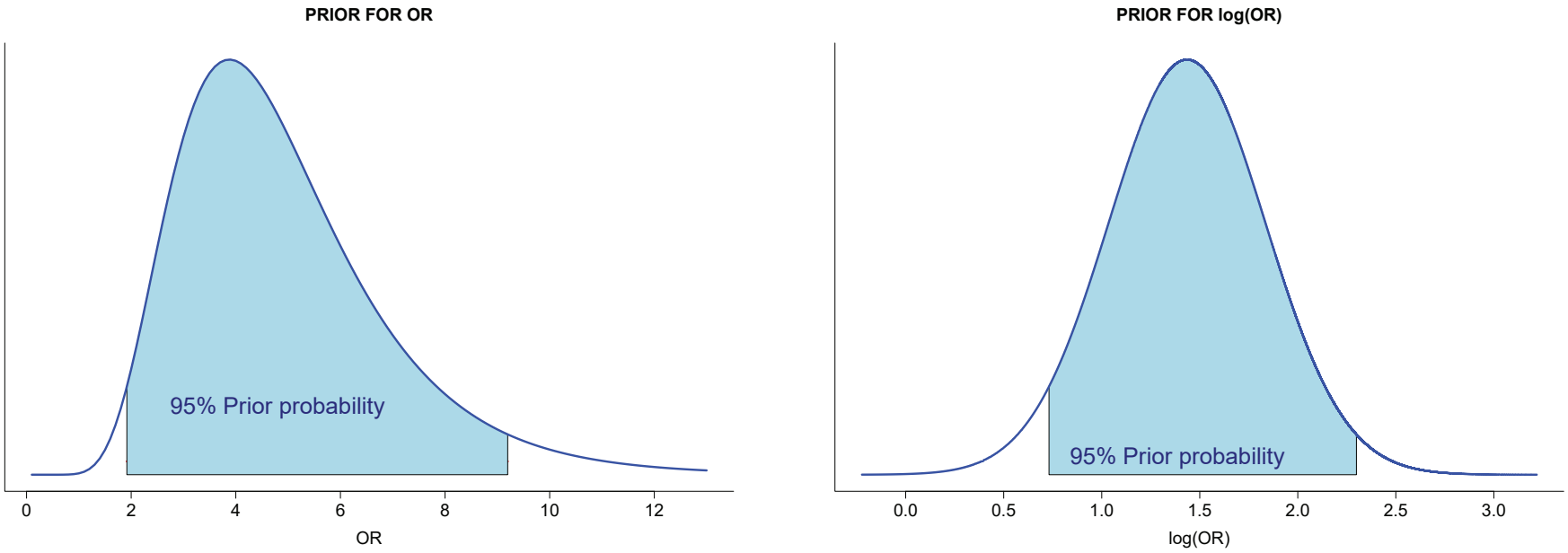
Turn information into a prior distribution



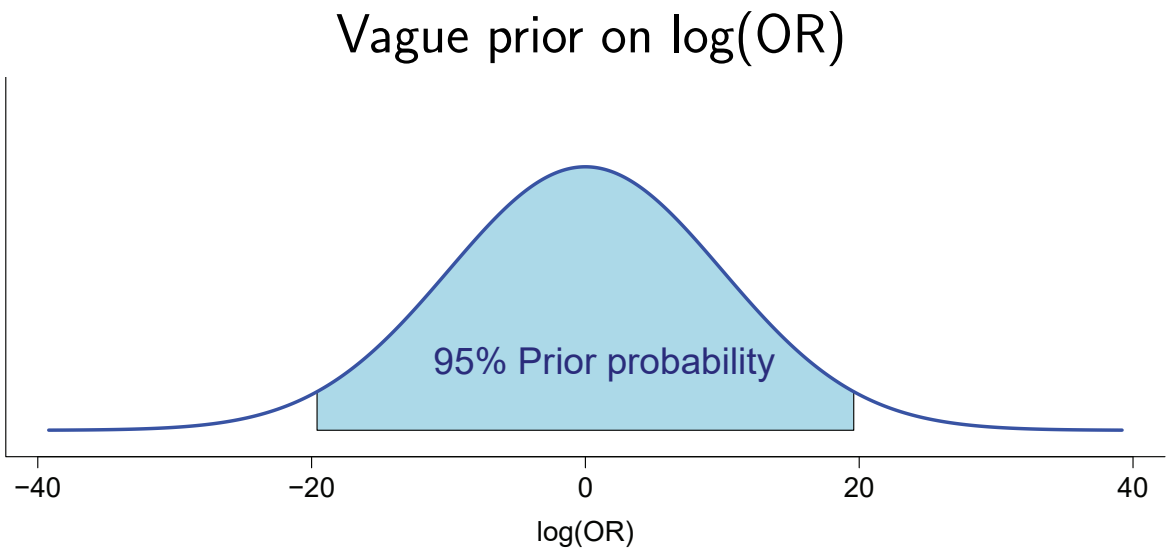
How to specify prior information on θ?

- Prior information can also be specified on log(OR)-scale
- ... and back transformed on OR-scale

Prior on OR and log(OR) from first 3 studies on SIDS



- One can also specify that there is no prior information
- ... as one does in a classical/frequentist approach
- Then one speaks of a non-informative prior, but better is to call it a vague prior

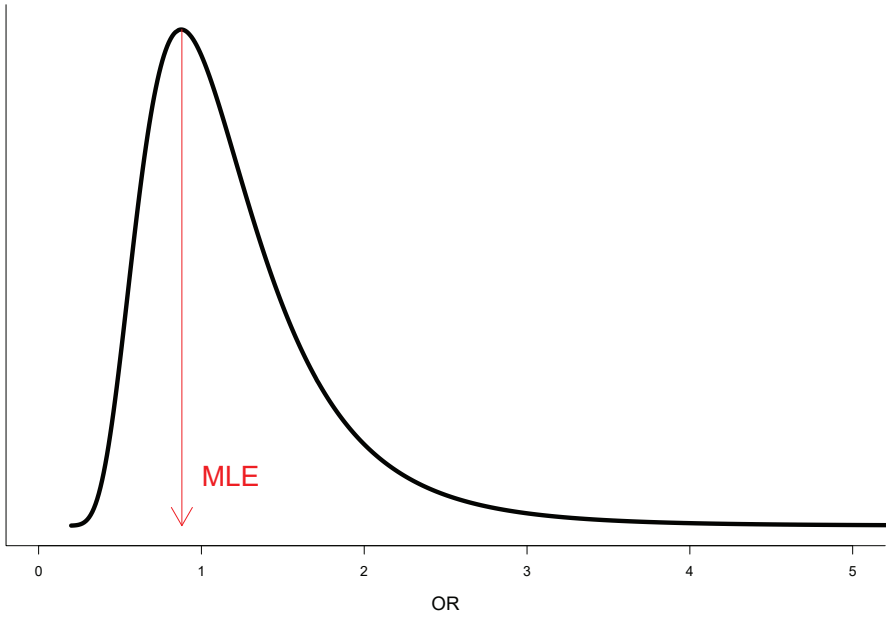


Note: This prior is in fact too "vague",
since one puts (very small) prior probability on unrealistic ORs

What information do the data bring on θ?

- Study results from Tonkin (1986):
  $\widehat{OR} = 0.88$ with 95% CI $=$ [0.40, 1.90] and P $= 0.74$

Likelihood function for θ



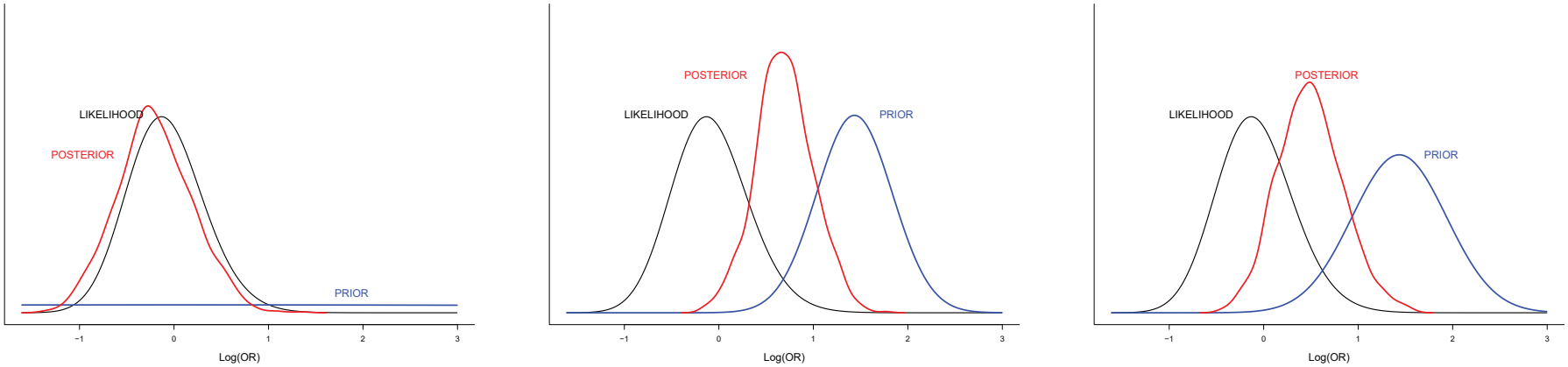
How to combine information on θ?

- Take prior on θ
- And likelihood on θ
- Take product of two functions for each possible value of θ
- Standardize $\Rightarrow$ Posterior (distribution)

$=$ Bayes Theorem

Prior, likelihood and posterior on log(OR)

Vague prior **Prior from 1st 3 studies** **Discounted prior**



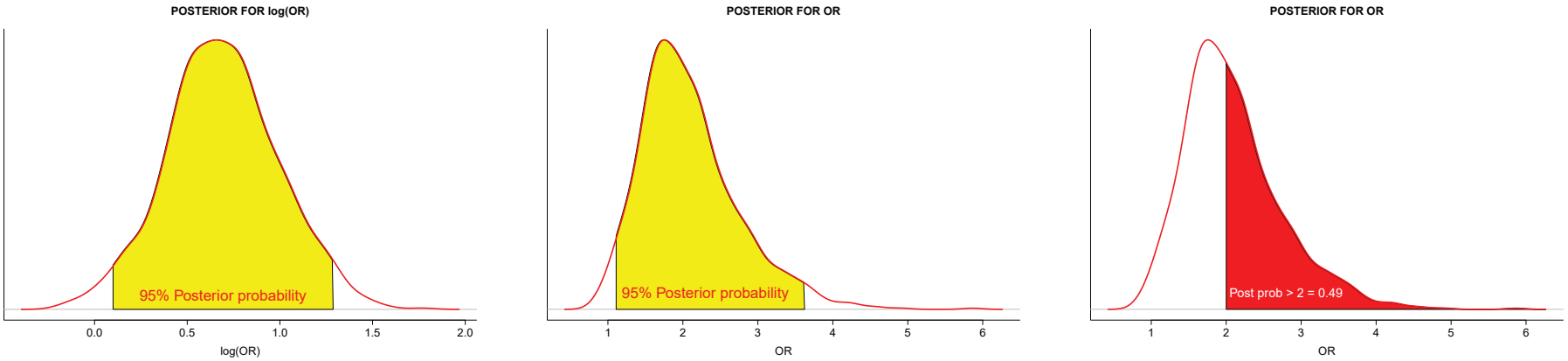
▷ Note:

- Prior, likelihood and posterior provide for each log(OR) its plausibility
- Posterior combines information from prior and likelihood
- **Posterior is a compromise between prior and likelihood**

▷ Figures can also be produced for odds-ratio

- **Posterior provides all information what researcher needs**

Posterior on OR and log(OR)



- Posterior probability $\Pr(\text{OR} > 1 \mid \mathcal{D})$ is easier to understand than P-value
- Also most probable value of OR, posterior mean & median & SD
- And ... Bayesian 95% confidence interval + ...

An often posed question

Is a Bayesian analysis just a cumulative meta-analysis?

Answer: The Bayesian approach offers much ... much more

- ▷ Prior knowledge can come from everywhere
- ▷ Bayesian approach can be applied to basically every problem
- ▷ Bayesian output, e.g. $\Pr(\text{OR} > 1 \mid \mathcal{D})$
- ▷ Bayesian software can handle problems that the classical approach cannot!

1.3 Bayes theorem – Formal definition

- Bayes Theorem combines prior and data information
  - ▷ Notation:
    - ◦ Unknown parameter log(OR) $= \theta$ & data: $\mathcal{D}$
    - ◦ Likelihood $L(\theta \mid \mathcal{D})$: plausibility of θ given data $\mathcal{D}$ for grid of θ values
    - ◦ Prior $p(\theta)$: prior density of θ values (information on θ independent of $\mathcal{D}$)
    - ◦ Posterior $p(\theta \mid \mathcal{D})$: posterior density of θ values as a result of combining prior and data information
  - ▷ General Bayes Theorem:

$$p(\theta \mid \mathcal{D}) = \frac{L(\theta \mid \mathcal{D})p(\theta)}{p(\mathcal{D})} = \frac{L(\theta \mid \mathcal{D})p(\theta)}{\int L(\theta|\mathcal{D})p(\theta)d\theta}$$

- **Thus**:
  - ▷ Posterior is the standardized **product of likelihood and prior**
  - ▷ Prior and posterior attach (Bayesian) probabilities to the different (intervals of) θ values
  - ▷ Thereby, unknown parameter gets a distribution, quite different from classical statistics

- Note:

More complex models

- Most statistical models involve more than one parameter, examples are:
  - Normal distribution: mean μ and variance σ^2
  - Linear regression: regression coefficients $\beta_0, \beta_1, \ldots, \beta_d$ and residual variance σ^2
  - Logistic regression: regression coefficients $\beta_0, \beta_1, \ldots, \beta_d$
  - Multinomial distribution: class probabilities $\theta_1, \theta_2, \ldots, \theta_d$ with $\sum_{j=1}^{d} \theta_j = 1$

- Bayes Theorem is the same for a single parameter as for multiple parameters:

$$p(\boldsymbol{\theta} \mid \mathcal{D}) = \frac{L(\boldsymbol{\theta} \mid \mathcal{D}) p(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta} \mid \mathcal{D}) p(\boldsymbol{\theta}) \, d\boldsymbol{\theta}}$$

  - Hence, the same expression as before but now $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_d)^\top$
  - Now, the prior $p(\boldsymbol{\theta})$ is multivariate. But often a prior is given for each parameter separately
  - Posterior $p(\boldsymbol{\theta} \mid \mathcal{D})$ is also multivariate. But we usually look only at each parameter separately, this involves the (marginal) posteriors $p(\theta_j \mid \mathcal{D})$ $(j = 1, \ldots, d)$

1.4 Bayesian computations in practice

- ▷ To apply Bayes Theorem, one needs to compute an often **intractable integral**
- ▷ **About 250 years** one has been looking how to do that in practice
- ▷ The solution lies in bypassing the integral by sampling the posterior, and thereby approximating the posterior summary measures
- ▷ This is what is called Markov chain Monte Carlo (MCMC) sampling
- ▷ Most famous general software is WinBUGS
- ▷ Nowadays also: OpenBUGS, JAGS, Stan, INLA + links to R/Python + dedicated R/Python packages

Instead of adapting the statistical model to the software

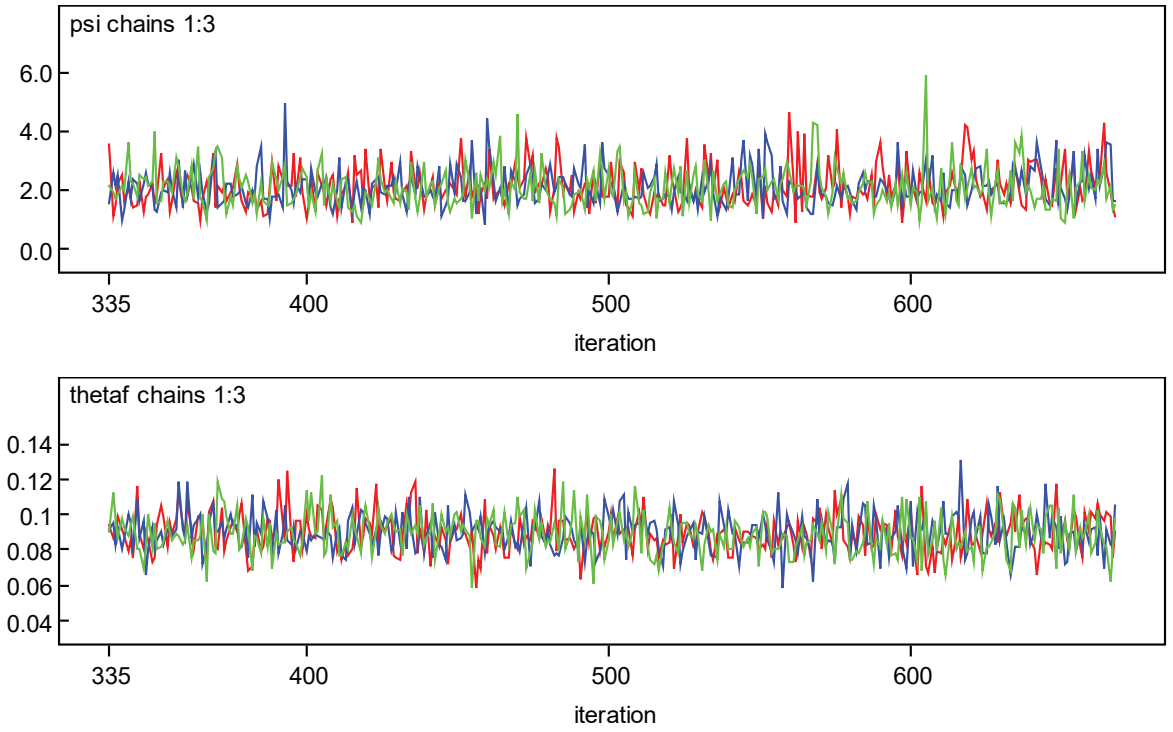
Bayesian software allows to swiftly
adapt the program to the statistical model

BUGS program: Infant sleeping position and SIDS

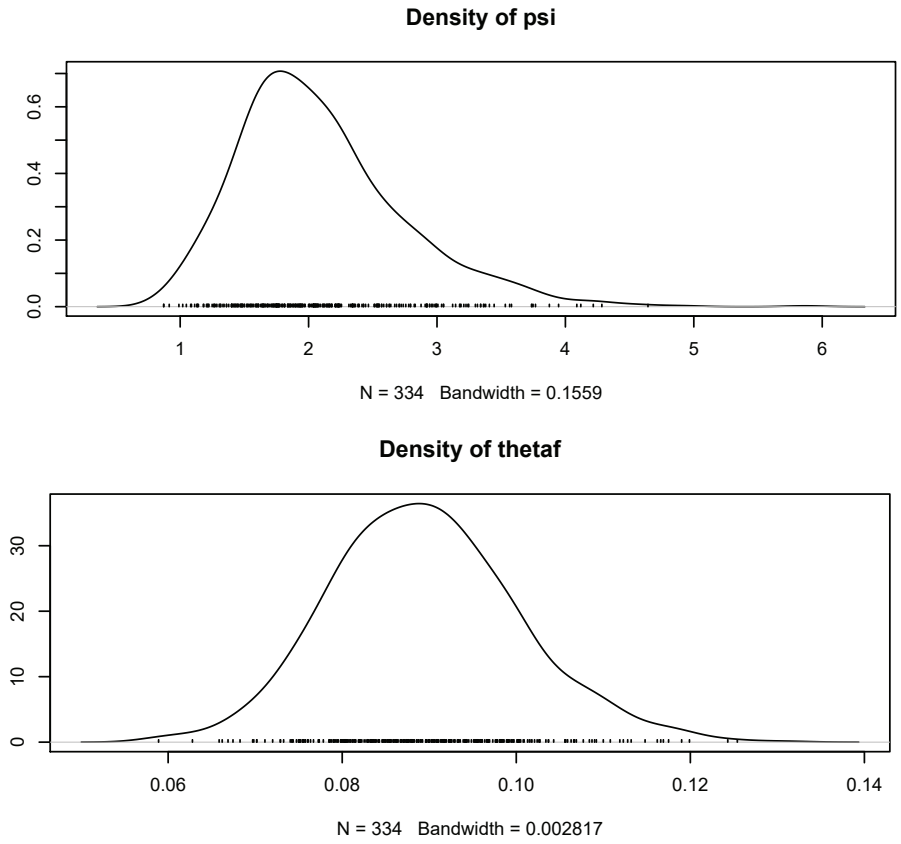
- MCMC computations were needed in the SIDS example
- WinBUGS program to compute posterior:

```
model {
rf ~ dbin(thetaf,nf)
rb ~ dbin(thetab,nb)
prec <- pow(sd,-2)
lpsi ~ dnorm(meanlogor,prec)
psi <- exp(lpsi)
thetaf  <- thetab*psi/((1-thetab)+thetab*psi)
thetab ~ dunif(0,1)
}
```

- **Trace plots**:


psi chains 1:3
6.0
4.0
2.0
0.0
335
400
500
600
iteration
thetaf chains 1:3
0.14
0.12
0.1
0.08
0.06
0.04
335
400
500
600
iteration

- **Posterior densities**:



- Summary measures:

```
1. Empirical mean and standard deviation for each variable,
plus standard error of the mean:

              Mean      SD  Naive SE Time-series SE
psi        2.09290 0.65384 0.0206555      0.0213636
lpsi       0.69277 0.30151 0.0095250      0.0098946
thetaf     0.08945 0.01087 0.0003435      0.0003678
thetab     0.04845 0.01432 0.0004525      0.0004778
deviance 14.43361 2.79413 0.0882698      0.0883546
```

1.5 Application to pharmaceutical research

▷ Suppose:

- one has some idea of the efficacy of a treatment versus control
- or ... one is skeptical that the new treatment is better than control
- or ... one does not have any idea of the efficacy of the new treatment

▷ and

- Upon completion of the study
- ... after having seen the results combined with prior belief
- we wish ... to assess the probability that the new treatment is better than control

▷ **Then: we need to use Bayes Theorem**

Why a Bayesian approach is useful or ...

may be even necessary in the design and analysis of RCTs

will be seen

after the next commercial spot

Session 2: The Bayesian approach in various research areas

Session 2
The Bayesian approach in various research areas

- The Bayesian approach to inference is increasing in popularity
- Applications can be found in:

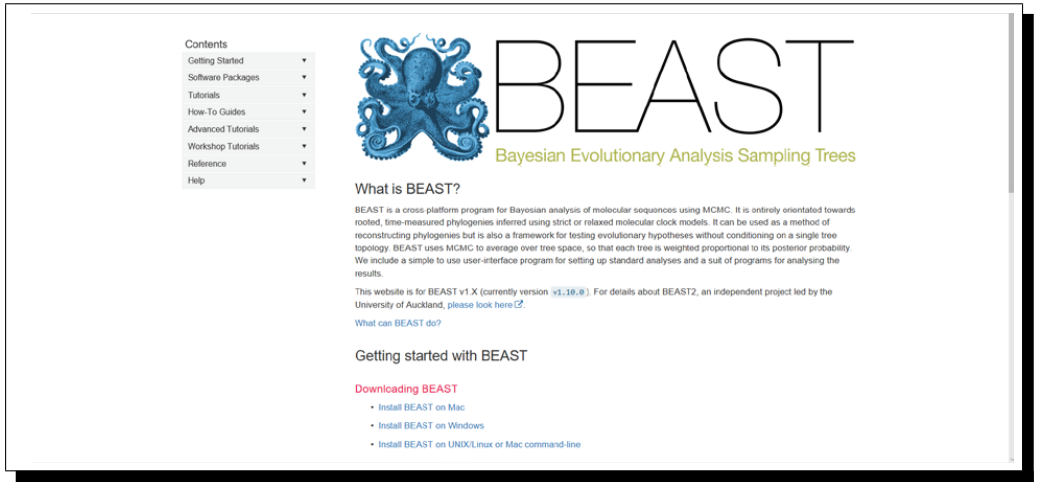
  Pharmaceutical research, Genetics, Epidemiology, Psychology, Sociology, Forensics, Archeology, Cosmology, Ecology, Marketing, Physics, ...

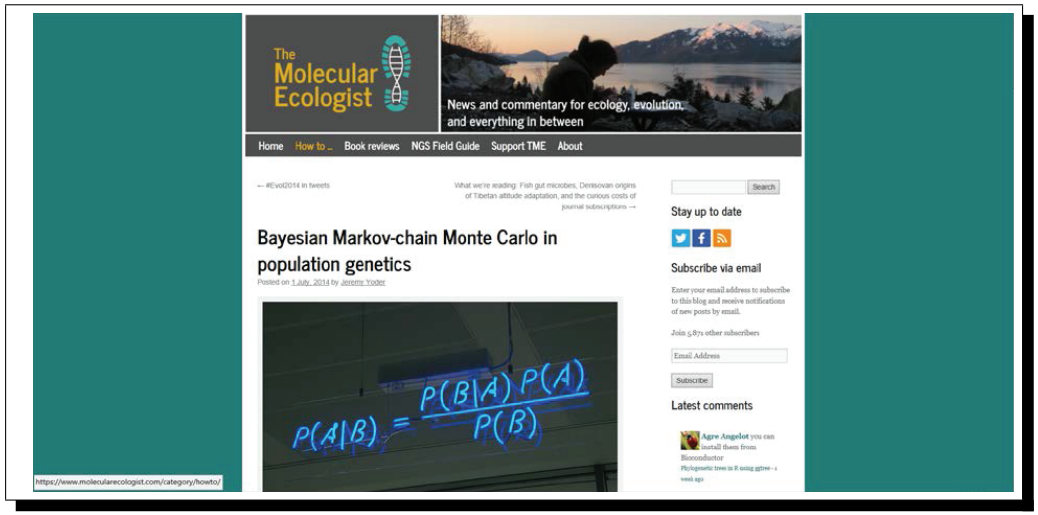
Pharmaceutical research

- Bayesian statistics in
  - Product development and manufacturing
  - Process development and validation
  - Analytical methods and assays
  - Stability studies
  - Sampling, release and content uniformity
  - Dissolution testing
  - Manufacturing
  - Medical devices
  - Program and portfolio decision-making

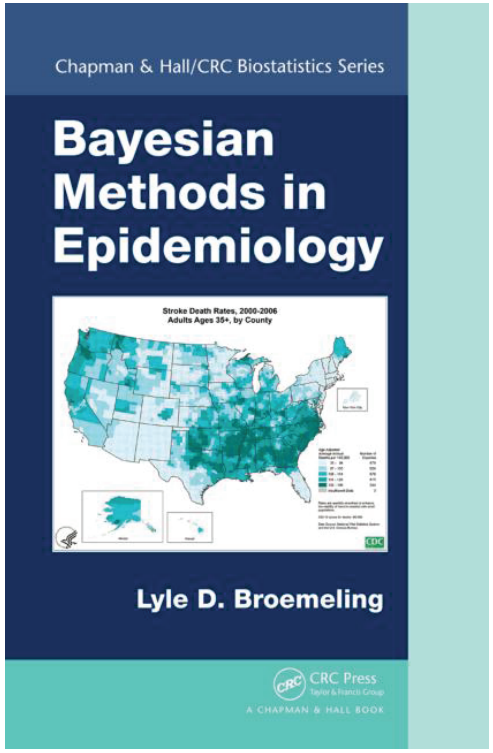
Bayes20XX Meetings

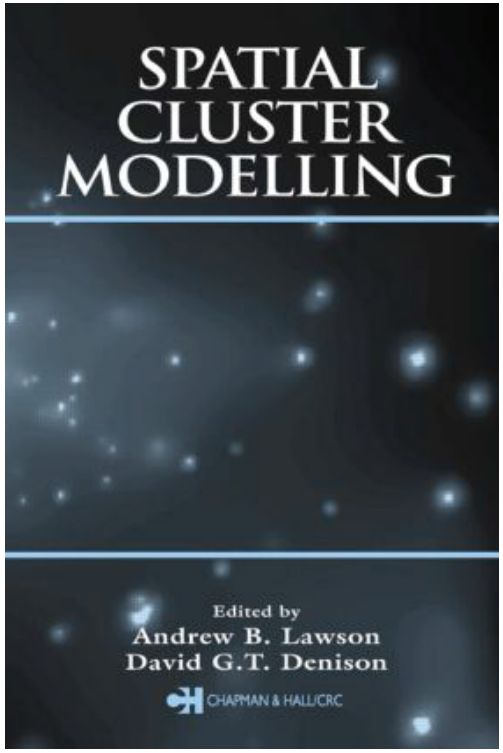
Genetics

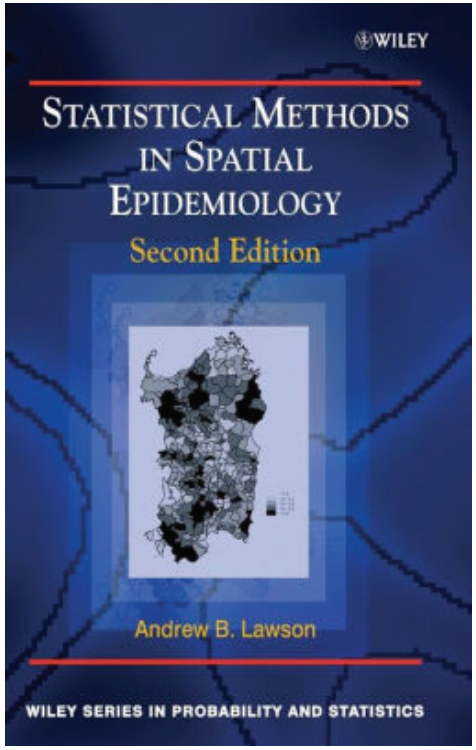

Contents
Getting Started
Software Packages
Tutorials
How-To Guides
Advanced Tutorials
Workshop Tutorials
Reference
Help
BEAST
Bayesian Evolutionary Analysis Sampling Trees
What is BEAST?
BEAST is a cross platform program for Bayesian analysis of molecular sequences using MCMC. It is entirely orientated towards rooted, time-measured phylogenies inferred using strict or relaxed molecular clock models. It can be used as a method of reconstructing phylogenies but is also a framework for testing evolutionary hypotheses without conditioning on a single tree topology. BEAST uses MCMC to average over tree space, so that each tree is weighted proportional to its posterior probability. We include a simple to use user-interface program for setting up standard analyses and a suit of programs for analysing the results.
This website is for BEAST v1.X (currently version v1.10.0). For details about BEAST2, an independent project led by the University of Auckland, please look here.
What can BEAST do?
Getting started with BEAST
Downloading BEAST
Install BEAST on Mac
Install BEAST on Windows
Install BEAST on UNIX/Linux or Mac command-line


The Molecular Ecologist
News and commentary for ecology, evolution, and everything in between
Home
How to ...
Book reviews
NGS Field Guide
Support TME
About
Bayesian Markov-chain Monte Carlo in population genetics
Stay up to date
Subscribe via email
Latest comments
https://www.molecularecologist.com/category/howto/

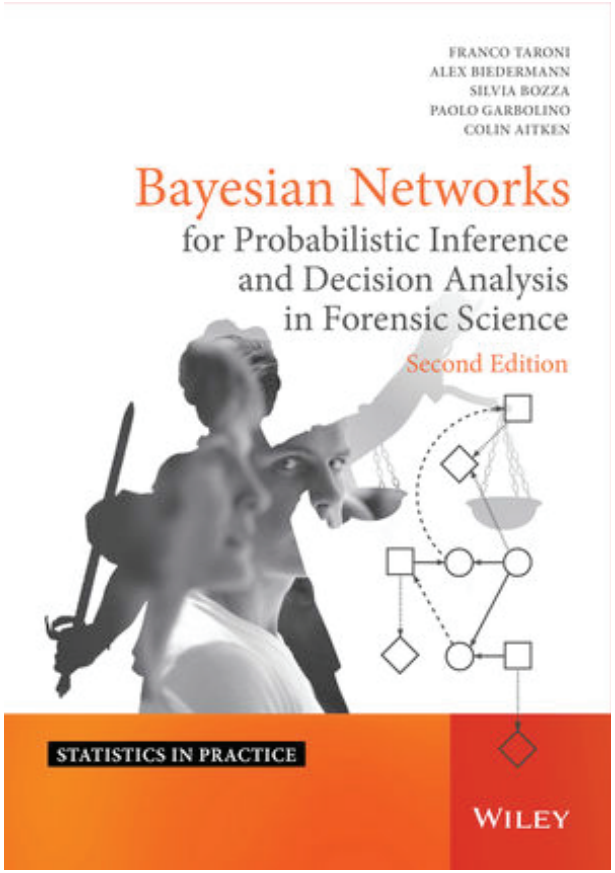
Epidemiology

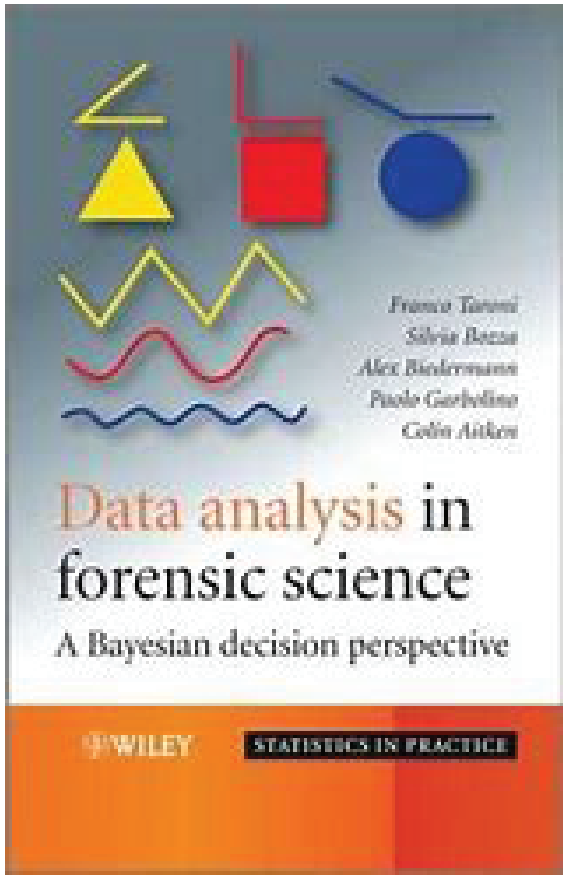






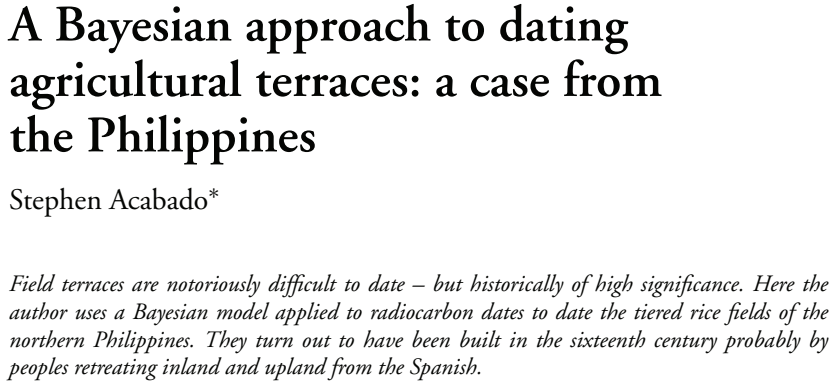
Forensics


FRANCO TARONI
ALEX BIEDERMANN
SILVIA BOZZA
PAOLO GARBOLINO
COLIN AITKEN
Bayesian Networks
for Probabilistic Inference
and Decision Analysis
in Forensic Science
Second Edition
STATISTICS IN PRACTICE
WILEY


Franco Taroni
Silvia Bozza
Alex Biedermann
Paolo Garbolino
Colin Aitken
Data analysis in
forensic science
A Bayesian decision perspective
WILEY
STATISTICS IN PRACTICE

Archeology

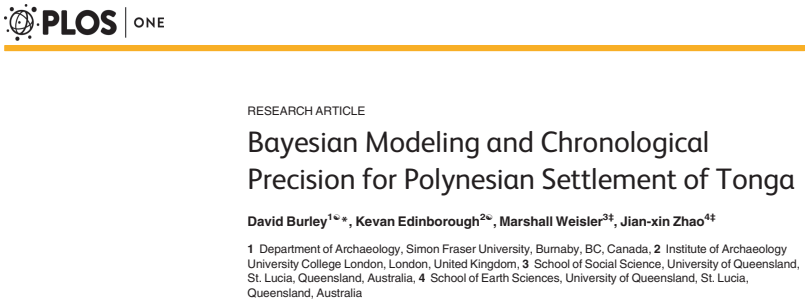
- An important activity in archeology is dating (no not dating ...) of found excavations
- Since data are usually not abundant in such activity, the Bayesian approach proves to be useful



A Bayesian approach to dating agricultural terraces: a case from the Philippines

Stephen Acabado*

Field terraces are notoriously difficult to date – but historically of high significance. Here the author uses a Bayesian model applied to radiocarbon dates to date the tiered rice fields of the northern Philippines. They turn out to have been built in the sixteenth century probably by peoples retreating inland and upland from the Spanish.



PLOS ONE

RESEARCH ARTICLE

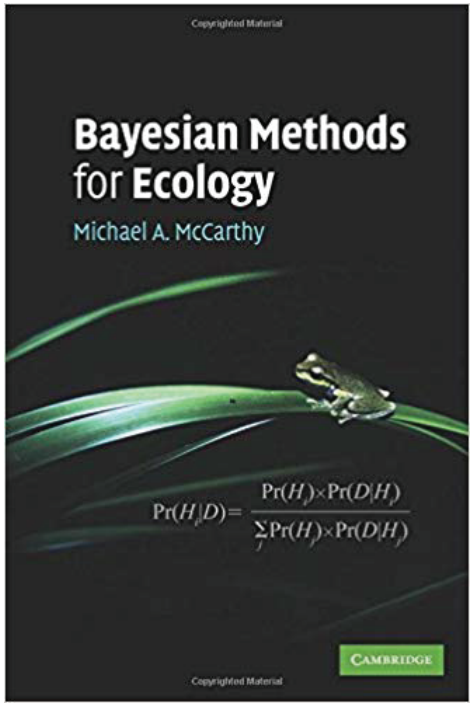
Bayesian Modeling and Chronological Precision for Polynesian Settlement of Tonga

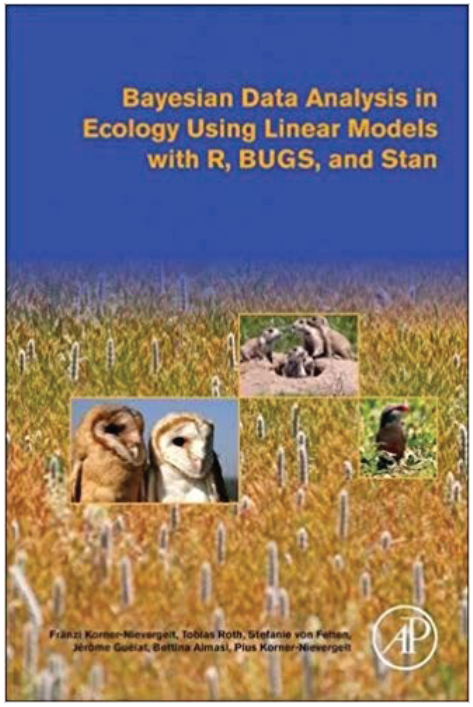
David Burley[1☯*], Kevan Edinborough[2☯], Marshall Weisler[3‡], Jian-xin Zhao[4‡]

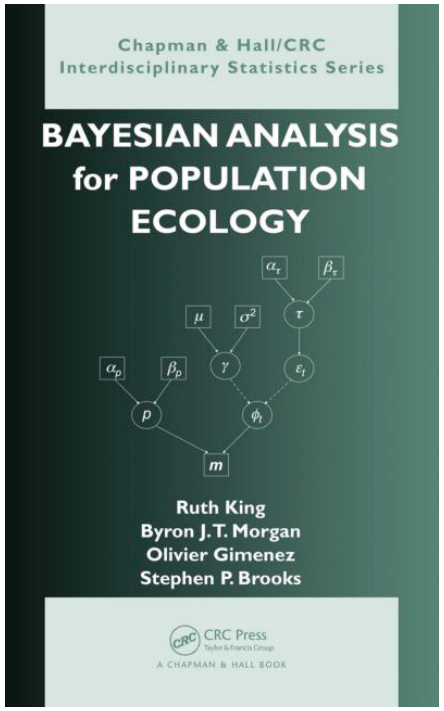
1 Department of Archaeology, Simon Fraser University, Burnaby, BC, Canada, 2 Institute of Archaeology University College London, London, United Kingdom, 3 School of Social Science, University of Queensland, St. Lucia, Queensland, Australia, 4 School of Earth Sciences, University of Queensland, St. Lucia, Queensland, Australia

Ecology

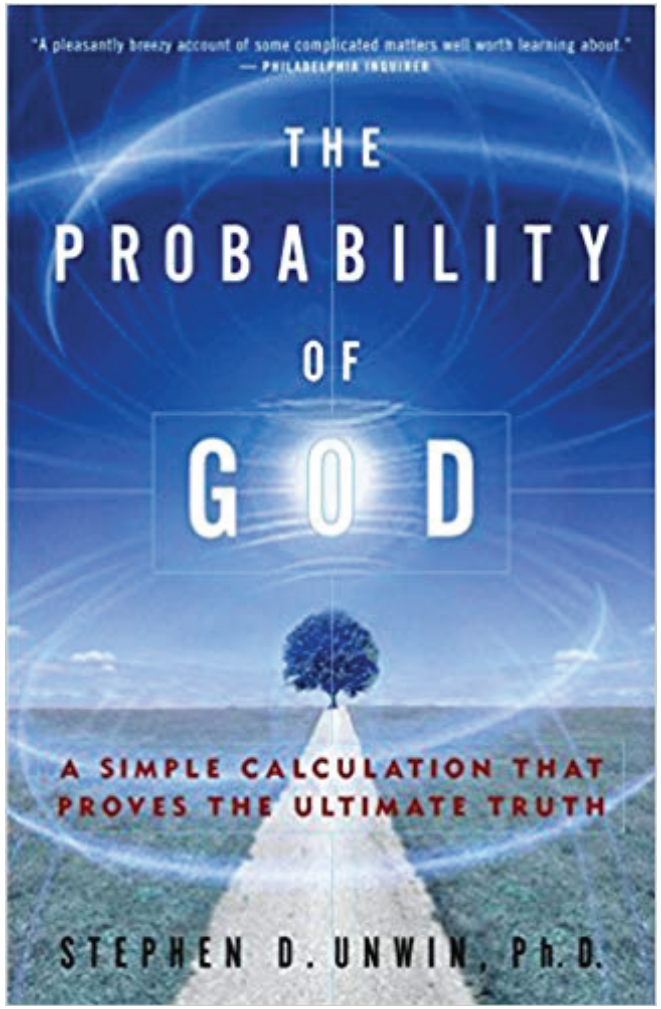
▷ Also in ecology, the Bayesian approach is becoming increasingly popular







Even in theology



And many many many other application areas ...

Session 3: Bayesian clinical trials

Session 3
Bayesian clinical trials

Here we illustrate the Bayesian approach for the design and analysis of RCTs:

- Simple Bayesian analyses of an RCT
- Bayesian approaches in modelling
- Bayesian interim analyses in phase II and phase III RCTs
- Bayesian approaches in dose finding studies
- Bayesian approaches for RCTs on orphan diseases
- Use of historical or additional information in analyzing RCTs
- Bayesian extrapolation studies
- Bayesian methods for medical device RCTs
- Bayesian adaptive RCTs

- **The traditional approach to statistics can still (and always) be used if**:
  - ▷ There is enough budget
  - ▷ Recruitment of patients goes fast
  - ▷ A control group can easily be taken
  - ▷ It is ethical to set up a classical RCT on children, pregnant women, ...
  - ▷ We don't worry about reinventing the wheel ...
- **But, what to do if that is not the case????**

3.1 Introduction

- The frequentist approach is still the most used approach in the setup, conduct and analysis of clinical trials

- However, the Bayesian approach is slowly but steadily gaining popularity:
  - ▷ In phase I trials (dose finding/safety trials) the Bayesian approach is one of the important tools
  - ▷ In phase II trials (proof-of-concept trials) the Bayesian approach is the preferred method in some of the major pharmaceutical companies
  - ▷ In medical device trials the Bayesian approach is one of the standard tools
  - ▷ But, in phase III trials the Bayesian approach is only used in interim analyses, but things are changing

3.2 Simple Bayesian analyses of an RCT

▷ In this section we show two Bayesian analyses:

1. **ARISTOTLE study**: large RCT originally analyzed in a frequentist way
   - Comparison of apixaban versus warfarin in patients with atrial fibrillation, Granger et al (NEJM, 2011)
   - Apixaban: vitamin K antagonist developed to prevent stroke in patients with atrial fibrillation (AF)
   - Apixaban is also called a novel oral coagulant (NOAC)

2. **SARS-CoV-2 (Pfizer) study**: large RCT with primary Bayesian efficacy analysis
   - Large RCT to evaluate safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine, Polack et al (NEJM, 2020)

ARISTOTLE study: characteristics

- ▷ DB RCT to compare apixaban (5 mg/bid) with warfarin (target INR 2.0 – 3.0)
- ▷ 18,201 (A: 9120, W: 9081) patients were recruited with AF + ≥ 1 additional risk factor for stroke
- ▷ **Primary outcome**: ischemic or hemorrhagic stroke or systemic embolism
- ▷ **Primary aim**: test for noninferiority, with threshold 1.38 for hazard ratio of apixaban/warfarin ($\text{HR}_{A/W}$)
- ▷ **Secondary objectives**: testing for superiority for primary outcome + rates of major bleeding and death from any cause

- ◦ Non-inferiority test: test whether apixaban is slightly worse than warfarin
- ◦ Superiority test: test whether apixaban is better than warfarin

ARISTOTLE study: results

- ▷ Median duration $=$ 1.8 years
- ▷ Rates primary outcome:
  - ◦ Apixaban: 1.27%/year
  - ◦ Warfarin: 1.60%/year
- ▷ $\text{HR}_{A/W}$: 0.75 with 95% CI $=$ [0.66, 0.95]
- ▷ 1P_noninferiority < 0.001, 2P_superiority < 0.01

ARISTOTLE study: Bayesian setting

- Assume here that all patients were followed-up for 2 years, and that we are interested in the proportion of patients that experience the primary event
- Let:
  - π_A = unknown proportion for apixaban
  - π_W = unknown proportion for warfarin
  - $\Delta = \pi_W - \pi_A$ = unknown efficacy for apixaban versus warfarin
  - 1.38 = NI threshold for HR $\approx \delta_{NI}$ = -0.01 = NI threshold for Δ
- Non-inferiority test (at $\Leftrightarrow$ 1-sided 0.05 significant):

$$\Pr(\Delta > \delta_{NI} \mid data) \geq 0.95$$

- Superiority test (at $\Leftrightarrow$ 2-sided 0.05 significant):

$$\Pr(\Delta > 0 \mid data) \geq 0.975$$

ARISTOTLE study: prior information

▷ We consider three cases:

Case 1. For neither π_A and π_W **we have prior knowledge** (**vague prior**)

Case 2. Warfarin was standard drug for AF patients ⇒ **informative prior for** π_W: around 3% with 95% uncertainty between 2% and 4%.

Case 3. **Skeptical prior for** Δ expressing that it will be probably close to zero: around 0% with 95% uncertainty between -2% and 2%

▷ Two study sizes are considered:

1. Each arm has 1000 patients $n_A = 1000$ and $n_W = 1000$
2. Apixaban $= 9120$ patients, Warfarin $= 9081$ patients

Bayesian analyses of the ARISTOTLE study: results

▷ Obtained proportions: $\widehat{\pi}_W = 0.029$ (2.9%) and $\widehat{\pi}_A = 0.023$ (2.3%)

▷ $\delta_{NI} = -0.01$

▷ Study sizes: $n_A = 1000$ and $n_W = 1000$

Vague NI test: $\Pr(\Delta > \delta_{NI} \mid data) = 0.997$, Sup test: $\Pr(\Delta > 0 \mid data) = 0.80$

Infor NI test: $\Pr(\Delta > \delta_{NI} \mid data) = 0.996$, Sup test: $\Pr(\Delta > 0 \mid data) = 0.86$

Skept NI test: $\Pr(\Delta > \delta_{NI} \mid data) = 0.991$, Sup test: $\Pr(\Delta > 0 \mid data) = 0.75$

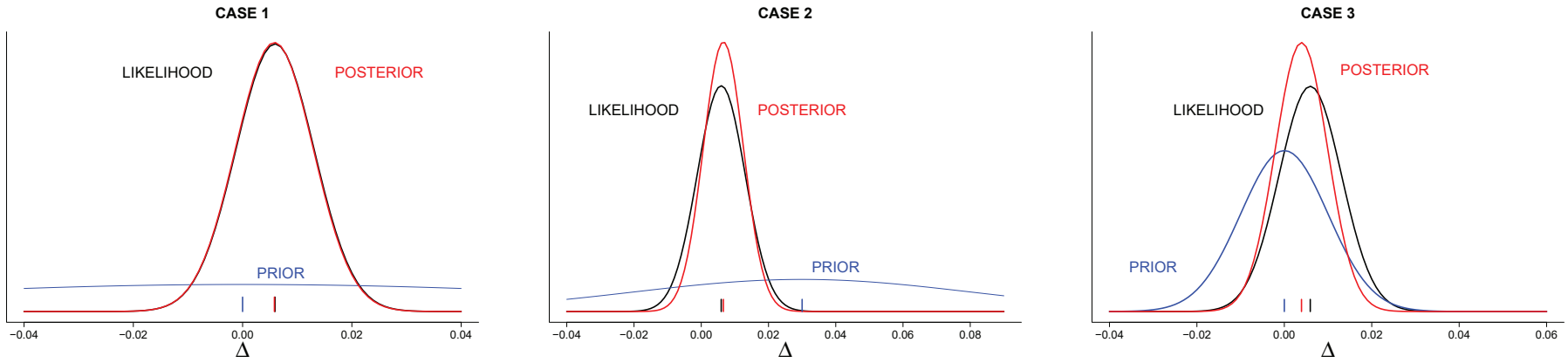
▷ Study sizes: $n_A = 9120$ and $n_W = 9081$

Vague NI test: $\Pr(\Delta > \delta_{NI} \mid data) = 1$, Sup test: $\Pr(\Delta > 0 \mid data) = 0.994$

Infor NI test: $\Pr(\Delta > \delta_{NI} \mid data) = 1$, Sup test: $\Pr(\Delta > 0 \mid data) = 0.996$

Skept NI test: $\Pr(\Delta > \delta_{NI} \mid data) = 1$, Sup test: $\Pr(\Delta > 0 \mid data) = 0.993$

ARISTOTLE study: graphical representation of Bayesian results



- **Vague (Case 1)**: posterior = likelihood ⇒ **frequentist result = Bayesian result**
- **Informative (Case 2)**: posterior takes historical information into account
- **Skeptical (Case 3)**: posterior takes skeptical opinion into account

SARS-CoV-2 (Pfizer) study

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 DECEMBER 31, 2020 VOL. 383 NO. 27

Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine

Fernando P. Polack, M.D., Stephen J. Thomas, M.D., Nicholas Kitchin, M.D., Judith Absalon, M.D.,

BACKGROUND

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and the resulting coronavirus disease 2019 (Covid-19) have afflicted tens of millions of people in a worldwide pandemic. Safe and effective vaccines are needed urgently.

METHODS

In an ongoing multinational, placebo-controlled, observer-blinded, pivotal efficacy trial, we randomly assigned persons 16 years of age or older in a 1:1 ratio to receive two doses, 21 days apart, of either placebo or the BNT162b2 vaccine candidate (30 μg per dose). BNT162b2 is a lipid nanoparticle–formulated, nucleoside-modified RNA vaccine that encodes a prefusion stabilized, membrane-anchored SARS-CoV-2 full-length spike protein. The primary end points were efficacy of the vaccine against laboratory-confirmed Covid-19 and safety.

RESULTS

A total of 43,548 participants underwent randomization, of whom 43,448 received injections: 21,720 with BNT162b2 and 21,728 with placebo. There were 8 cases of Covid-19 with onset at least 7 days after the second dose among participants assigned to receive BNT162b2 and 162 cases among those assigned to placebo;

Statistical analysis

Analysis of the first primary efficacy end point included participants who received the vaccine or placebo as randomly assigned, had no evidence of infection within 7 days after the second dose, and had no major protocol deviations (the population that could be evaluated). Vaccine efficacy was estimated by $100 \times (1 - \text{IRR})$, where IRR is the calculated ratio of confirmed cases of Covid-19 illness per 1000 person-years of follow-up in the active vaccine group to the corresponding illness rate in the placebo group. The 95.0% credible interval for vaccine efficacy and the probability of vaccine efficacy greater than 30% were calculated with the use of a Bayesian beta-binomial model. The final analysis uses a success boundary of 98.6% for probability of vaccine efficacy greater than 30% to compensate for the interim analysis and to control the overall type 1 error rate at 2.5%.

3.3 Bayesian approaches for statistical modelling

- The Bayesian approach is useful for complex data structures
- But also to allow for non-standard distributions or problems
- This is partly due to flexibility of Bayesian software (MCMC approach)
- For instance, the Bayesian approach is useful when:
  - response has a non-standard distribution, e.g. a t-distribution, skew-normal distribution, etc.
  - there are missing data in the response and/or covariates (possibly of a complex nature)
  - some model parameters are constrained
- **Examples**:
  - joint modelling of longitudinal and survival outcomes
  - longitudinal models of mixture distributions

3.4 Bayesian interim analyses in phase II/III RCTs

- ▷ The Bayesian approach is increasingly used to monitor clinical trials
- ▷ In interim analyses one:
  - ◦ verifies safety issues
  - ◦ checks for early stopping for efficacy
  - ◦ checks for early stopping futility analysis
- ▷ Examples illustrate the use of the Bayesian approach for interim analyses
- ▷ In the Bayesian approach there is **no correction for multiple views**, stopping only depends on:
  - ◦ the posterior probability of a treatment effect
  - ◦ the likely effect at the end obtained from the posterior predictive distribution

Bayesian (interim) analyses in a COVID-19 study

- COMPILE study:
  - ▷ A recently finished study (Goldfeld et al., 2021) coordinated by NYU to evaluate effect of convalescent plasma (CP) on COVID-19 patients
  - ▷ Example of meta-analytic approach to pooling individual patient data from completed, early-terminated and ongoing 8 RCTs for CP
  - ▷ Note: the CP arm is always the same, but the control groups differ
  - ▷ Multi-country/multi-centre study utilizing continuous monitoring using Bayesian stopping rules motivated by:
    - ◦ Frequentist stopping rules are far too complicated
    - ◦ Use of posterior probability as tool for monitoring is quite intuitive
    - ◦ Computer simulations establish control of Type I error rate
  - ▷ Primary endpoint(s): (1) WHO clinical status at 14 $+/-1$ days, (2) Mechanical ventilation or worse (1 vs 0)

WHO scale

TABLE 1 WHO 11-point COVID scale definition

0:	Uninfected, no viral RNA detected
1:	Asymptomatic, viral RNA detected
2:	Symptomatic, independent
3:	Symptomatic, assistance needed
4:	Hospitalized, no oxygen therapy
5:	Hospitalized, oxygen by mask or nasal prongs
6:	Hospitalized, oxygen by non-invasive ventilation or high flow
7:	Intubation & mechanical ventilation, $pO_2/FiO_2 \geq 150$ (or $SpO_2/FiO_2 \geq 200$)[1]
8:	Mechanical ventilation, $pO_2/FiO_2 < 150$ (or $SpO_2/FiO_2 < 200$) or vasopressors
9:	Mechanical ventilation, $pO_2/FiO_2 < 150$ and vasopressors, dialysis, or ECMO[2]
10:	Dead

[1] pO_2: partial pressure of oxygen, FiO_2: fraction of inspired oxygen, SpO_2: oxygen saturation

[2] ECMO: extracorporeal membrane oxygenation

Primary statistical analyses

A Bayesian analysis for all outcomes (primary+secondary+exploratory+safety)

- ▷ Outcome 1 (or): Ordinal random effects (trial) logistic regression analysis
- ▷ Outcome 2 (bin): Binary random effects (trial) logistic regression analysis
- ▷ In all analyses confounders (age, sex, ..) are taken in the models
- ▷ All model parameters have been given a prior distribution:
  - ◦ Most parameters have been given a vague prior, effect of CP is given a **skeptical prior**
  - ◦ Bayesian monitoring is based on posterior probabilities (previous example), **stop when**:

$$\Pr(OR_{or} < 1) \geq 0.95 \ \& \ \Pr(OR_{or} < 0.8) \geq 0.5$$

and

$$\Pr(OR_{bin} < 1) \geq 0.95 \ \& \ \Pr(OR_{bin} < 0.8) \geq 0.5$$

  - ◦ Priors based on extensive simulations so that Bayesian monitoring is compatible with frequentist monitoring

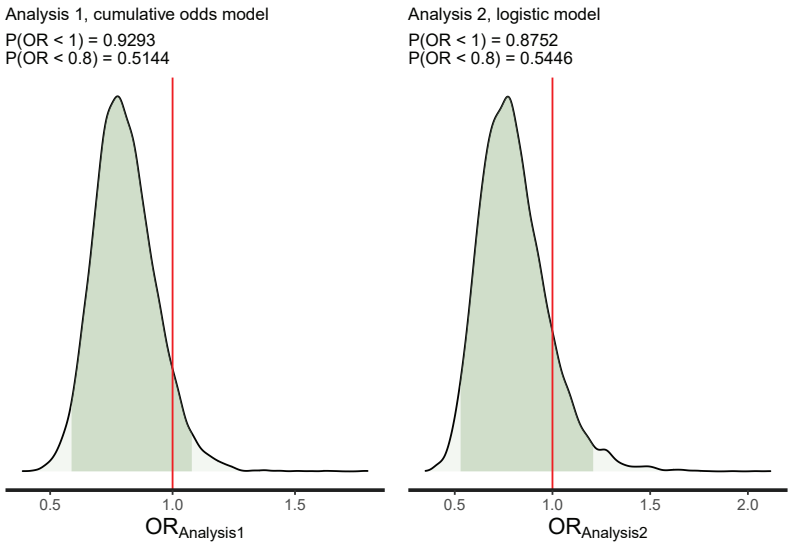
Bayesian interim analyses

- Bayesian interim analyses were organized at the following dates:
  - November 20, 2020
  - December 18, 2020
  - January 22, 2021
  - February 5, 2021
  - March 5, 2021
- The final analysis was done on end May 2021
- Unfortunately, the **criteria were not met** $\Rightarrow$ it could not be concluded that covalescent plasma was beneficial for all Covid-19 patients

Two interim analyses

Parsimonious adjustment includes *age*, *sex*, *WHO score at baseline*, and *days since symptom onset*.

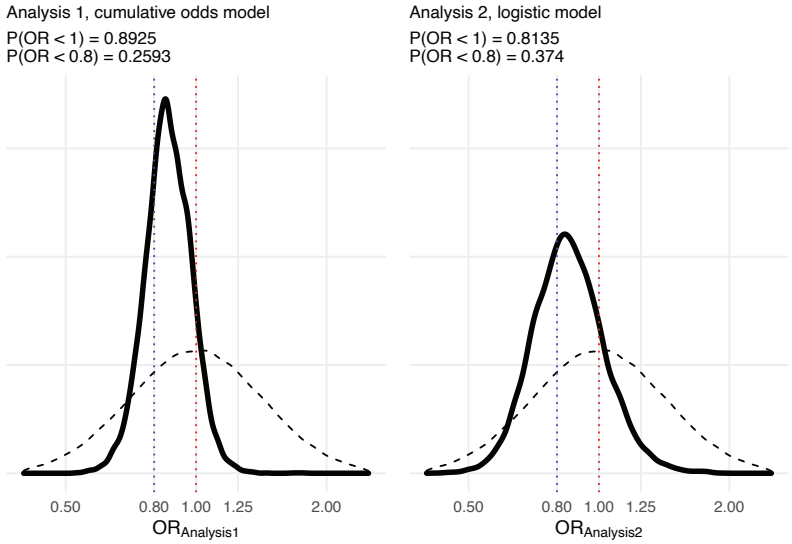
Stopping for efficacy: $P(OR < 1) \geq 0.95$ AND $P(OR < 0.8) \geq 0.5$ for both outcomes



November 11, 2020

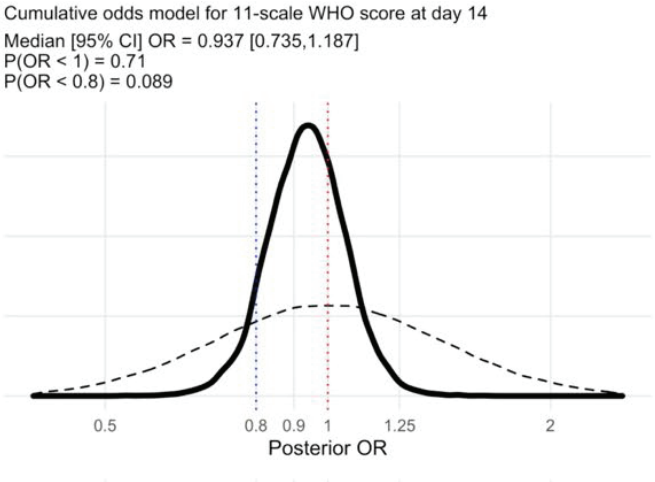
Parsimonious adjustment includes *age*, *sex*, *WHO score at baseline*, *days since symptom onset* and *enrollment quarters*.

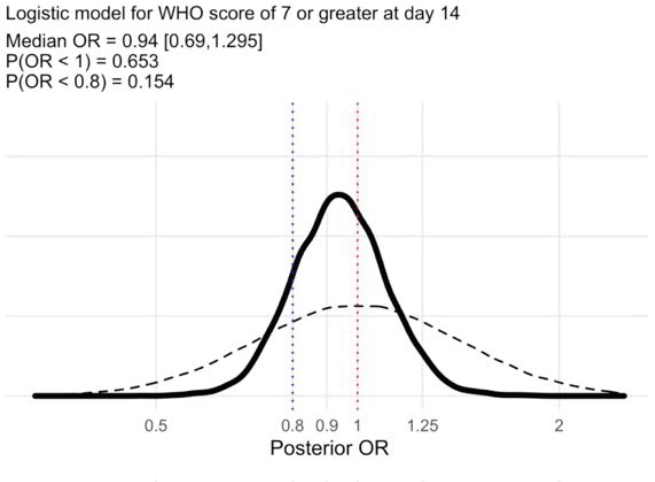
Stopping for efficacy: $P(OR < 1) \geq 0.95$ AND $P(OR < 0.8) \geq 0.5$ for both outcomes



January 22, 2021

Final analysis



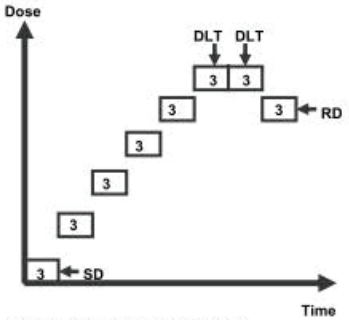


3.5 Dose-finding designs in phase I cancer studies

- Phase I dose escalation for cytotoxic agents against cancer
- Aim: Find the highest dose with acceptable toxicity
  - Optimal Dose = MTD = Maximum Tolerated Dose, for phase II
  - Toxicity used to define MTD: DLT = Dose Limiting Toxicities
- Characteristics phase I dose escalation studies:
  - Late stage patients
  - Small sized trials
  - Mix of tumor types
- Two main approaches:
  - Rule-based approach: **3 + 3 rule**
  - Model-based approach: **Bayesian modelling** (CRM & EWOC)

Two main approaches

▷ Rule-based approach: **3 + 3 rule**



from Le Tourneau (JNCI 2009)

▷ Model-based approach: **Bayesian modelling** (CRM & EWOC)

- Assume a monotonic relationship between dose and Pr(DLT)
- Dynamically adapt the dose at each subject based on previous observations
- Dose adaptation based on predicted toxicity at each possible dose
- Prediction based on a model for toxicity estimated from previous observations
- Unlike classical design: Probability of toxicity at a specific dose based on observations at all doses

3.6 Bayesian approaches for RCTs on orphan diseases

- ▷ A rare disease (orphan disease): any disease that affects a small percentage of the population
- ▷ Recently an inflation of rare diseases: due to sub-partitions of diseases
- ▷ Orphanet: 6,000 to 7,000 rare diseases, WHO: 5,000 to 8,000 rare diseases in 2013
- ▷ FDA: ≥ 600 drugs for orphan diseases
- ▷ Major difficulty with rare diseases: **huge sample sizes** are needed
- ▷ Use of prior information is an increasingly popular way to reduce sample size
- ▷ Here: example of a drug that was approved by regulatory agencies after a Bayesian analysis

Hemophilia safety study

- Hemophilia A is a genetic disease in which the clotting factor VIII is deficient with an increased bleeding tendency (and can be life threatening)

- Hemophilia (both A & B) are rare diseases–prevalence happens once in every 5,000 male births. Hemophilia B only happens in every 20,000 to 34,000 births.

- Treatment of disease:
  - Factor VIII (FVIII) from donated blood plasma, or DNA recombinant FVIII
  - Possible adverse events: transmission of diseases (blood donation) or antibody formation (**inhibitor**) to factor VIII (**DNA recombinant FVIII**)

- In 2005:
  - Refacto® AF was developed and one had to show its safety!
  - Known: 6 out 329 patients treated by **4 other similar compounds** developed inhibitor

FDA regulations

- Refacto® AF was based on, at that time, more sophisticated manufacturing processes
- The assessment of inhibitor risk in clinical trials of new or modified recombinant FVIII products is challenging, due to the low frequency of inhibitor occurrence and the generally small size of hemophilia studies due to the rarity of the disease
- **FDA** stated that for the trial to be successful, the upper bound of the two-sided (frequentist) 95% CI for the product inhibitor incidence rate must be below **6.8%**
- Sample size implications of the FDA requirement: unrealistic (between 7,000 and 50,000)
- Bayesian scenario was suggested and accepted by FDA

Bayesian approach

▷ **New (more stringent) Bayesian threshold**:

- Published ITT data on pivotal registration trials for 4 licensed products
- Combined data with Beta(1,1) prior, gives Beta(7,324) posterior
- This has 7/(7+324) = **2.1%** mean rate and **99%-ile = 4.4%**

⇒ **4.4%** was chosen as new threshold

▷ **Strategy for new study**:

- Inhibitor ITT data obtained from earlier studies with Refacto give discounted prior Beta(2.5,110)

▷ **Study on Refacto® AF**:

- New study with 90 patients
- 2 patients developed inhibitor out of 94 (ITT) patients
- Upper bound of 95% equal tail CI = **4.07%** < **4.4%**

▷ Source: Recht et al (2009)

3.7 Use of historical or additional information in RCTs

- There are numerous settings where historical or extra information is useful/can be used in the design and analysis of the current study:
  - Within a research institute an experiment has been repeated several times with the same control group
  - An experimental drug has been tested in many stages before going to phase III
  - An RCT is not possible because recruitment will be too slow, but extra information is available from either historical studies or from real world data
  - ...
- So it is either doing nothing or ... a Bayesian approach with an informative prior
- **Question**: How much of prior information can/should be used?

Approaches to include historical/additional information:

- ▷ **Bayesian borrowing (information) approaches**:
  - ◦ Bayesian basic principles: as in hemophilia example
  - ◦ Power prior: HOVON studies
  - ◦ Meta-analytic prior: HOVON studies & Transplant study
  - ◦ Commensurate prior: ...
- ▷ **Application areas**:
  - ◦ Orphan diseases
  - ◦ Pediatric studies
  - ◦ Bridging studies
- ▷ One can have prior information from a single historical study or from multiple sources

3.8 Bayesian methods for medical devices

- There are at least 4 reasons why the Bayesian approach is useful for the analysis of medical device trials:
  - Reason 1: Number of patients treated with medical devices is usually small
  - Reason 2: Process of innovation is rapid and continuously incremental
  - Reason 3: Endpoints may be far in the future
  - Reason 4: Same difficulties as for surgery, blinding is difficult
- Concurrent controls are often hard to take, historical information must often be used
- Therefore, Bayesian approach is an accepted approach by FDA, EMA
- FDA guidance document can be downloaded

3.9 Bayesian adaptive designs

- **Adaptive design**: Design is based on the results of the trial sofar
- Examples:
  - Two-stage design: re-estimation of sample size after seen interim data
  - Dose-finding design: next dose depends on results on previous selected doses
  - Adaptive randomisation: randomize next patient to treatment with highest posterior probability of efficacy
  - Seamless phase II-III trials: deleting or adding an arm to the trial, adapting dose, etc.
  - Master protocol trials: platform trials, umbrella trials, basket trials
- Aim:

To speed up pharmaceutical process

- Bayesian approach plays an important role in this!

Overview adaptive designs

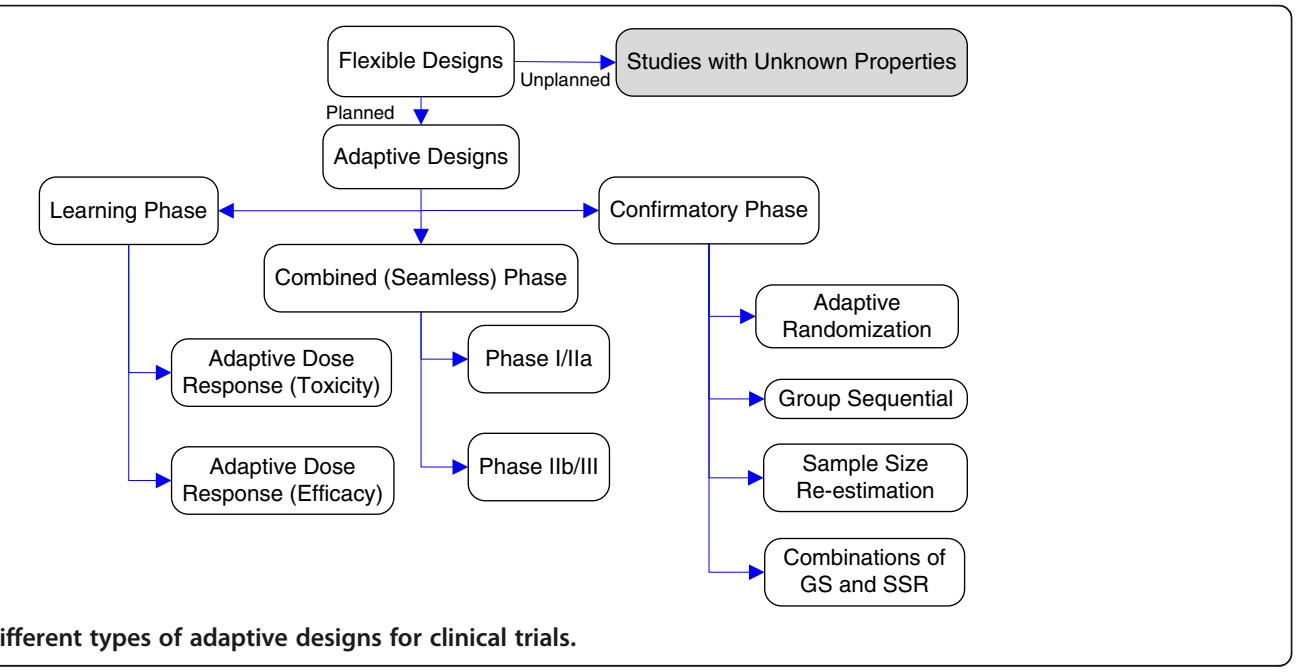


Figure 1 Summary of different types of adaptive designs for clinical trials.

Kairalla, Coffey, Thomann, Muller (TRIALS, 2012)

3.10 Platform trials, basket trials & umbrella trials

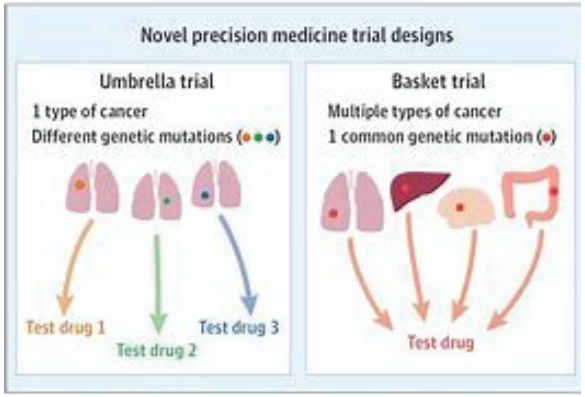
▷ Platform trial:

- finding the best treatment for a disease by simultaneously investigating multiple treatments, using specialized statistical tools for allocating patients and analyzing results
- Focus is on the disease rather than any particular experimental therapy

▷ For a good introductory reference: Berry, Connor, Lewis (JAMA, 2015)

▷ Platform trials, basked trials & umbrella trials make use of a **master protocol**:

- Master Protocol: An over-arching protocol or trial mechanism comprised of several parallel sub-trials differing by molecular features
- Example in oncology:
  - Basket Trial: A master protocol where each sub-trial enrolls multiple tumor types ("the basket")
  - Umbrella Trial: A master protocol where all patients (and all sub-trials) share a common tumor type ("the umbrella")



3.11 Additional topics

- Interesting and useful developments in pharmaceutical research have been made using Bayesian approach:
- Examples:
  - Estimation of accrual patients
  - Economic evaluations/Cost effectiveness
  - Product development and manufacturing
  - Pocess development and validation
  - Stability studies
  - Content uniformity studies
  - ...
- Check Lesaffre, Baio and Boulanger (2020) for more details

3.12 Obstacles for the Bayesian approach

- In the past: **no computational tools** to do a Bayesian analysis in practice
- Now: tradition, tradition and tradition
- But also:
  - Master programs in statistics have little training in Bayesian methods
  - The religious status of the P-value
  - Your ideas???
- But, things are improving:
  - FDA, EMA: there is more willingness to accept a Bayesian application

THE END

finally