Welcome in Louvain-la-Neuve

History of Louvain-la-Neuve (Université Catholique de Louvain)

1968

1968

1970

1972

1972

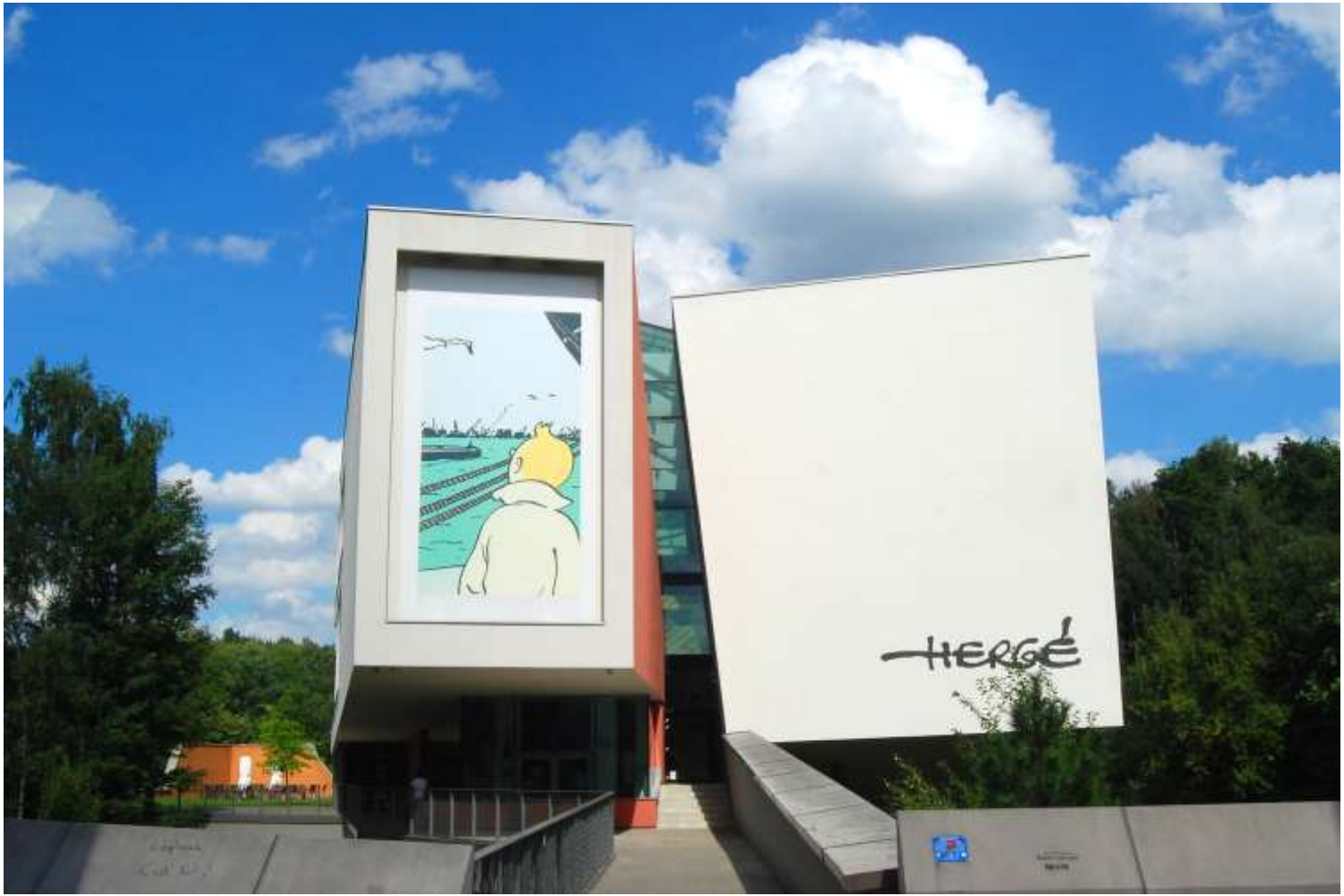
1984

1985

PHARMALEX

STATISTICAL SOLUTIONS

EFSPI



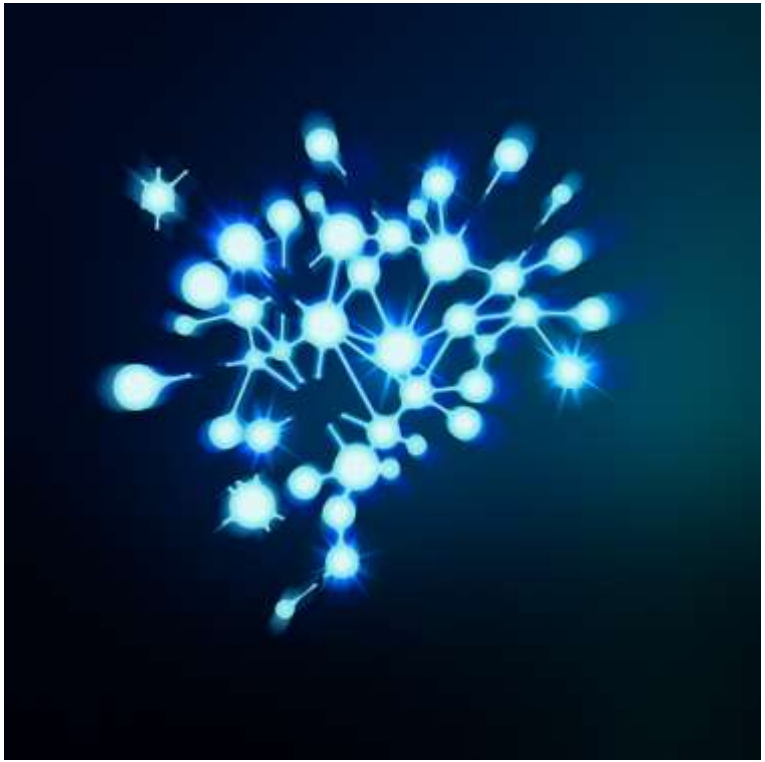
The raise of Bayesian statistics in pharmaceutical development
How to organize the transition?

Boulanger Bruno | CSO Arlenda

EFSPI Leadership Meeting
12 July 2018
Louvain-la-Neuve

Agenda

- **Survey**
- Drug Development as learning process
- The reproducibility crisis
- Bayesian reasoning
- How to organize the transition ?



Survey about Bayesian BioStatistics

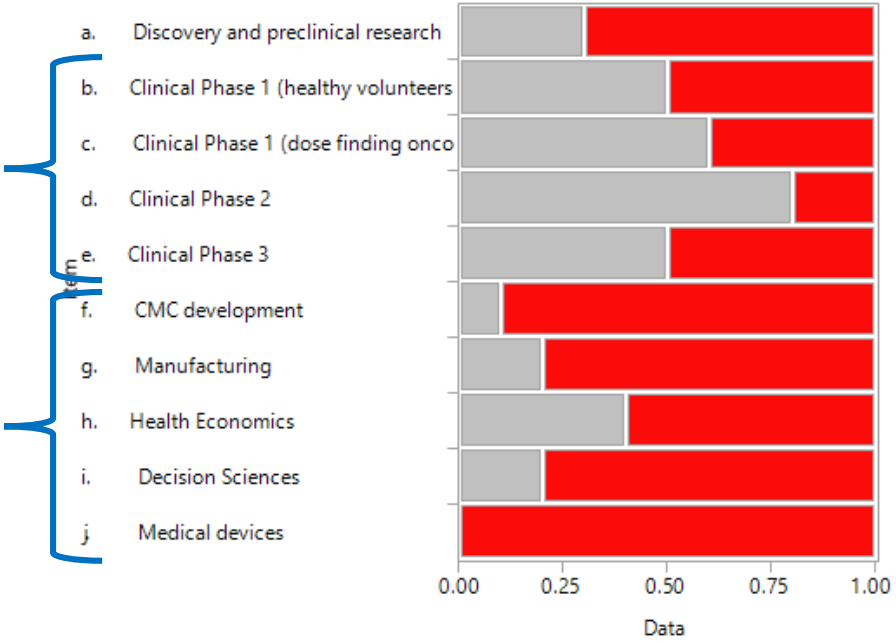
- Feedback from 11 companies
  - Abbott
  - Amgen
  - BI
  - Celgene
  - ChiesiFarma
  - Danone
  - IQVIA
  - Janssen
  - Roche
  - Servier
  - UCB

Summary results

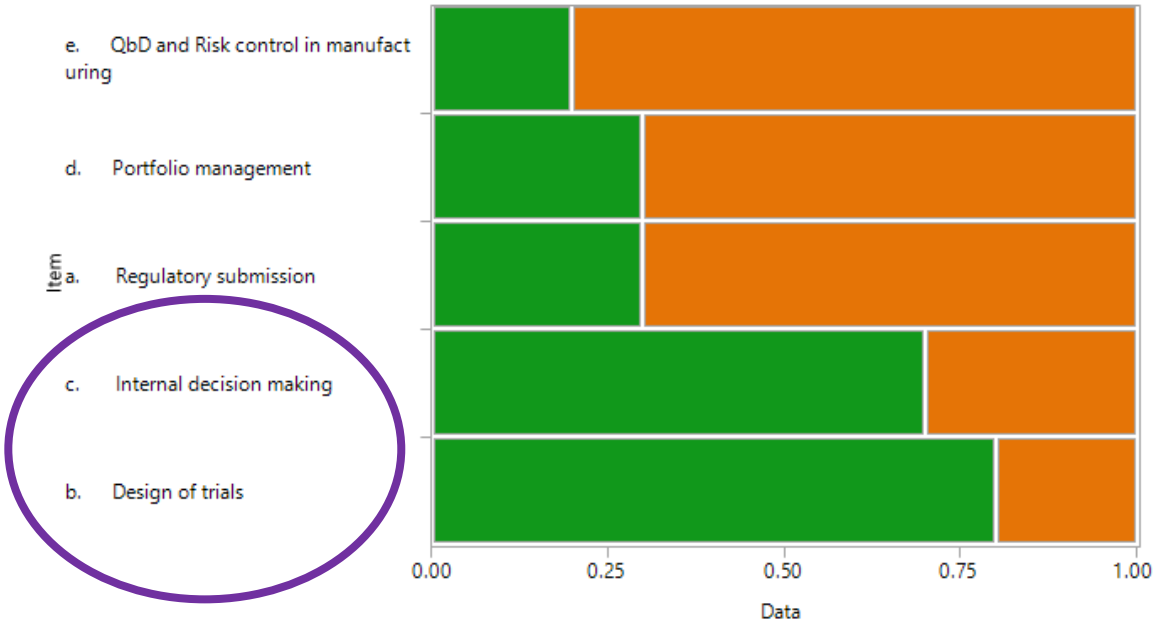
- 1- Do you currently apply elements Bayesian statistics within your company? **90%**
- 2- Phase do you use or intend to use Bayesian statistics?

Clinical phases are dominant in a challenging regulatory environment

There are great industrial opportunities with more regulatory flexibility

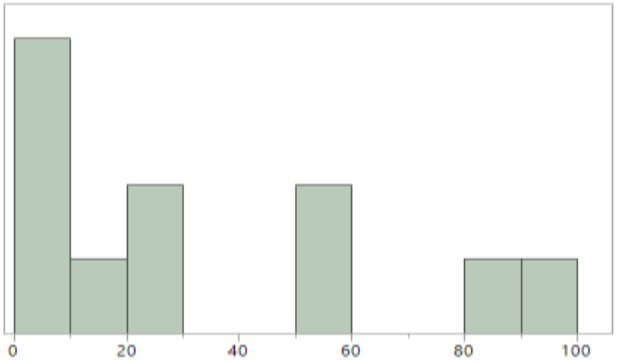


If you apply Bayesian statistics, is it for:

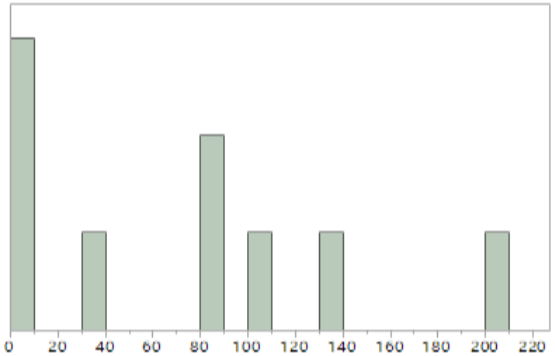


Training in Bayesian Biostatistics

- Percentage of trained statisticians



Number of statisticians



- Trained, aware, but maybe very little experience
- About 50% of statisticians have at least some knowledge. (334/648)

- 11 / 11 companies say they plan to train their statisticians
- 9 / 11 say they also plan to train the non-statisticians
- 3 / 11 already attended a dedicated conference in Bayesian BioStatistics

The main road blocks in use of Bayesian statistics

Sounds like a priority

Suggests more practical and applied trainings are needed

Suggests global perception is appropriate

The perceived acceptance by regulatory authorities

Statisticians are not trained in applying the concepts

The availability of programming languages /software tools

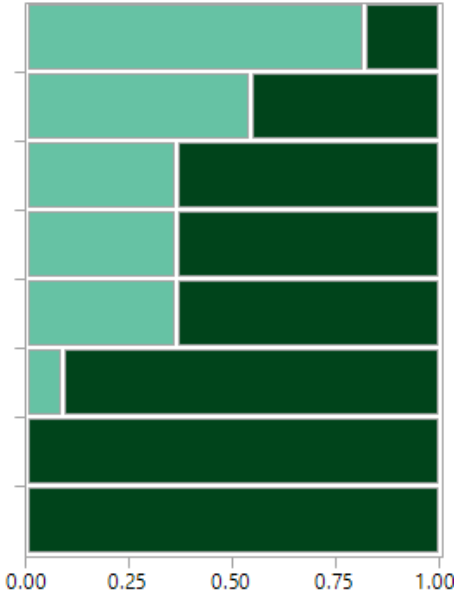
Lack of necessary capacities / too complex

Type I error is not controlled

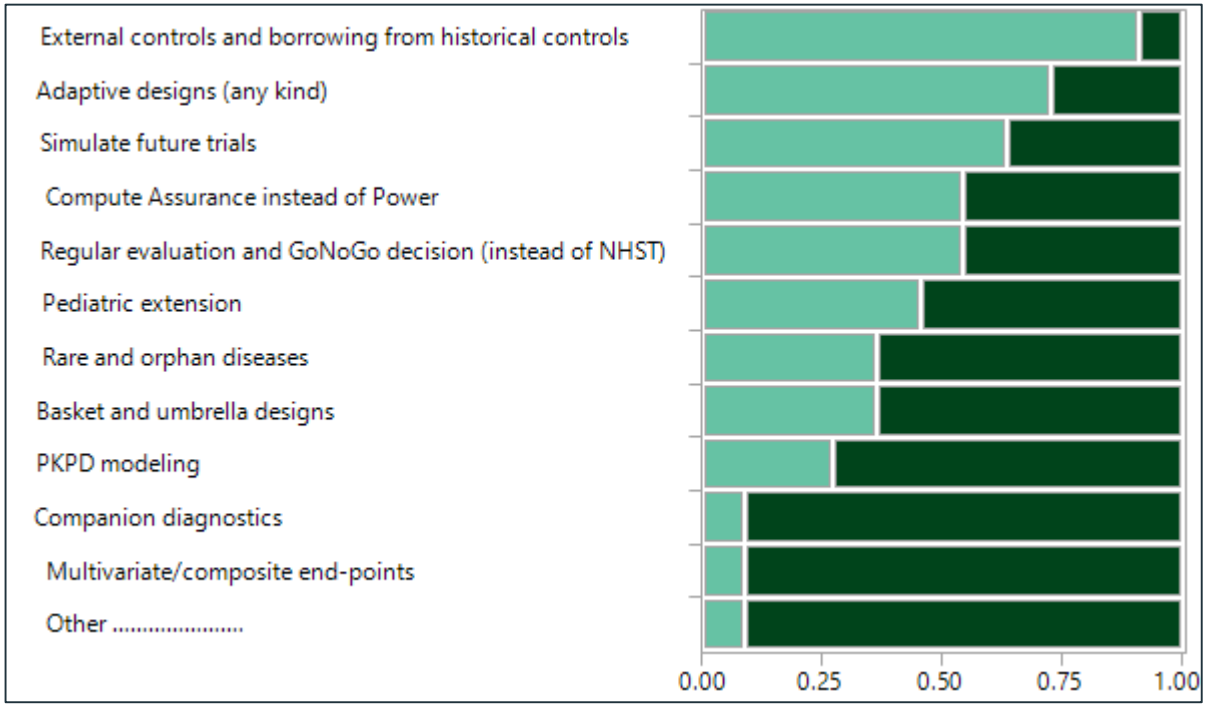
Difficulty to assess the convergence when MCMC methods are used

It doesn't bring added value compare to NHST

Estimates are biased



Bayesian statistics are used for

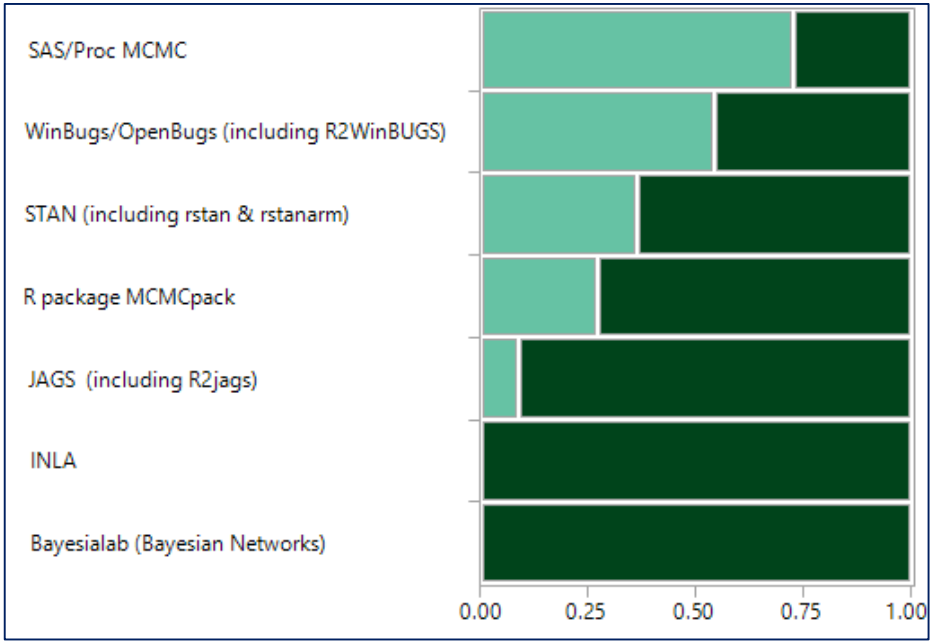


The greatest opportunities of Bayesian statistics

- The future is bright ;-)

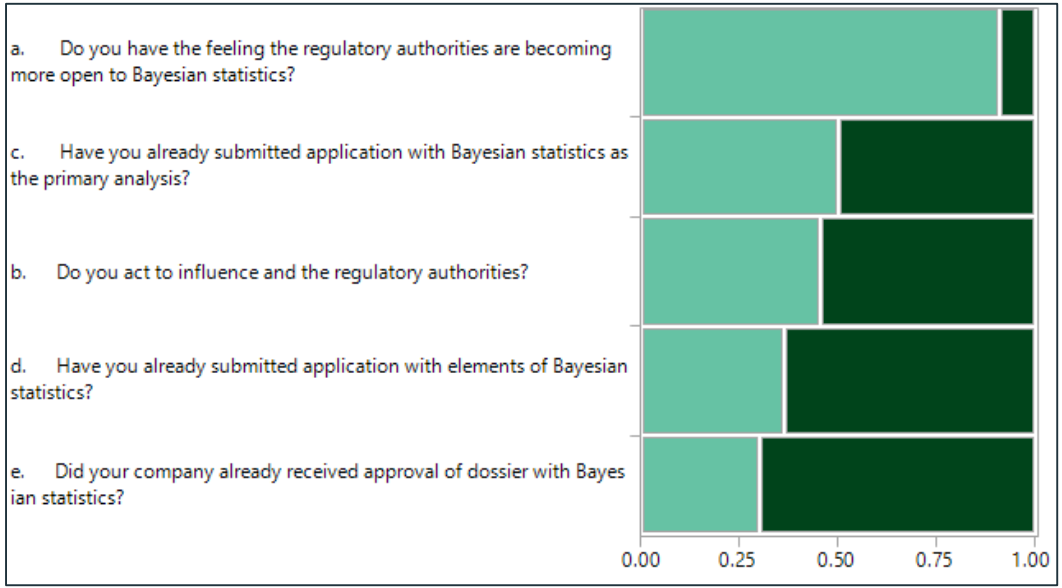
Most common Bayesian language

- WinBugs is still alive !
- SAS managed to reach first rank in few years
- STAN is already in the place



Regulatory acceptance

- The regulatory world is moving forward
- FDA is perceived as being ahead of EMA with that respect
- Companies keep the momentum



How to move to the next stage ?

Sounds like a priority. How ?

Collab. with SAS, STAN,…?

There is already one Group. Need more dedicated groups?

There is already such a conference

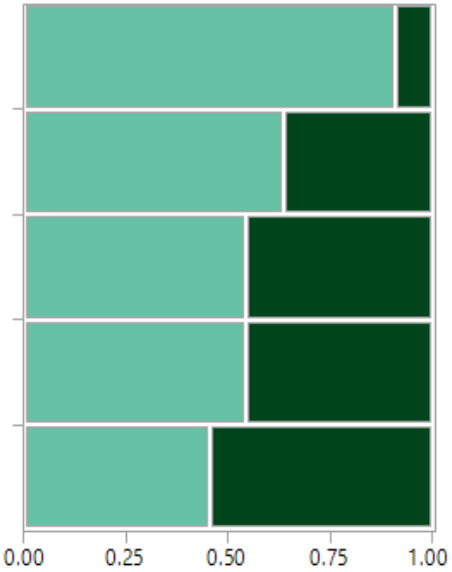
e. Initiate a platform of brainstorming with regulatory statisticians

a. Share examples via a web platform

d. Having a Working Group within the EFSPI dedicated to Bayesian statistics

b. Stimulate universities to increase the number of Bayesian statistics courses in the regular cursus

c. Organize conference in Applied Bayesian statistics



BAYES 2013
21-23 May ERASMUS UNIVERSITY ROTTERDAM

BAYES 2016
LEUVEN

BAYES2014

BAYES 2017
LA MANCHA
22-23-24 MAY 2017

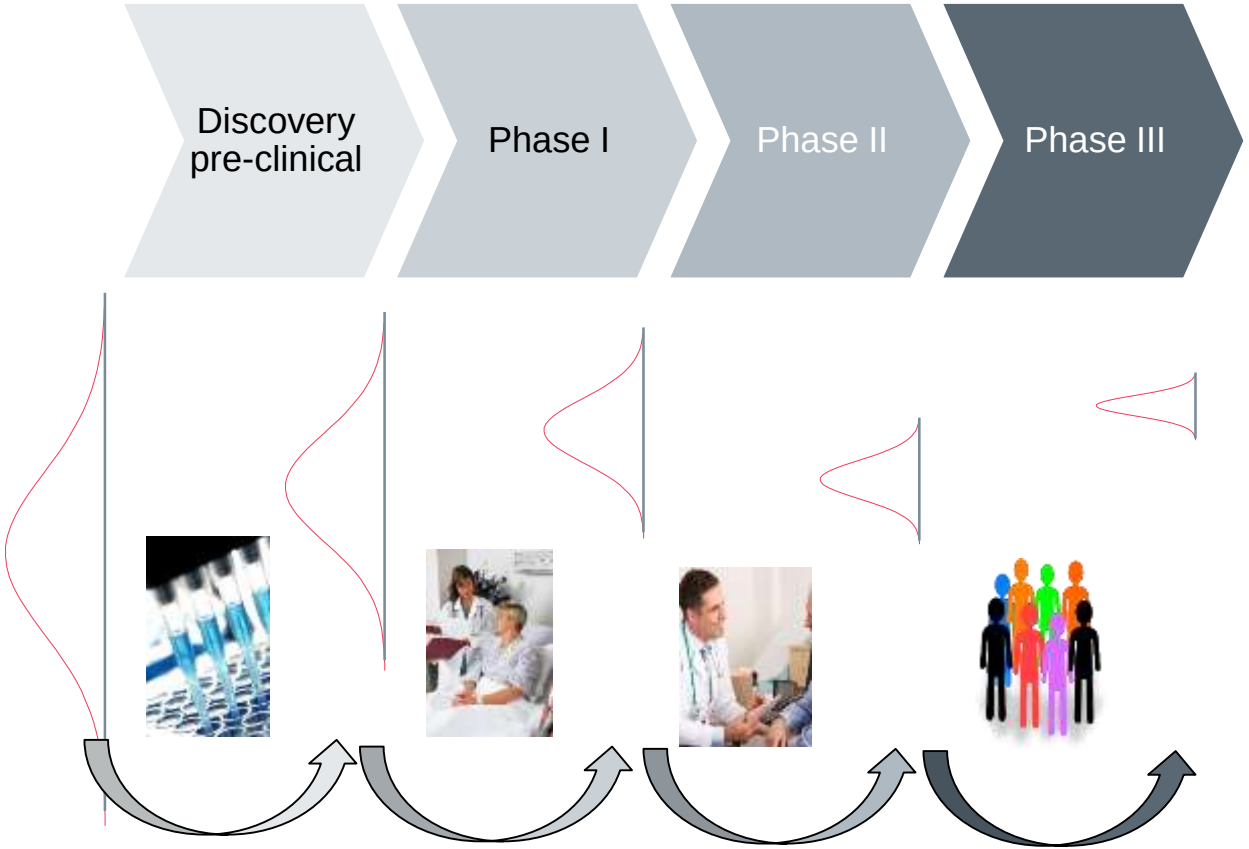
BAYES 2012
Aachen 9-10-11 MAY

BAYES2015
BASEL

BAYES 2018
CAMBRIDGE

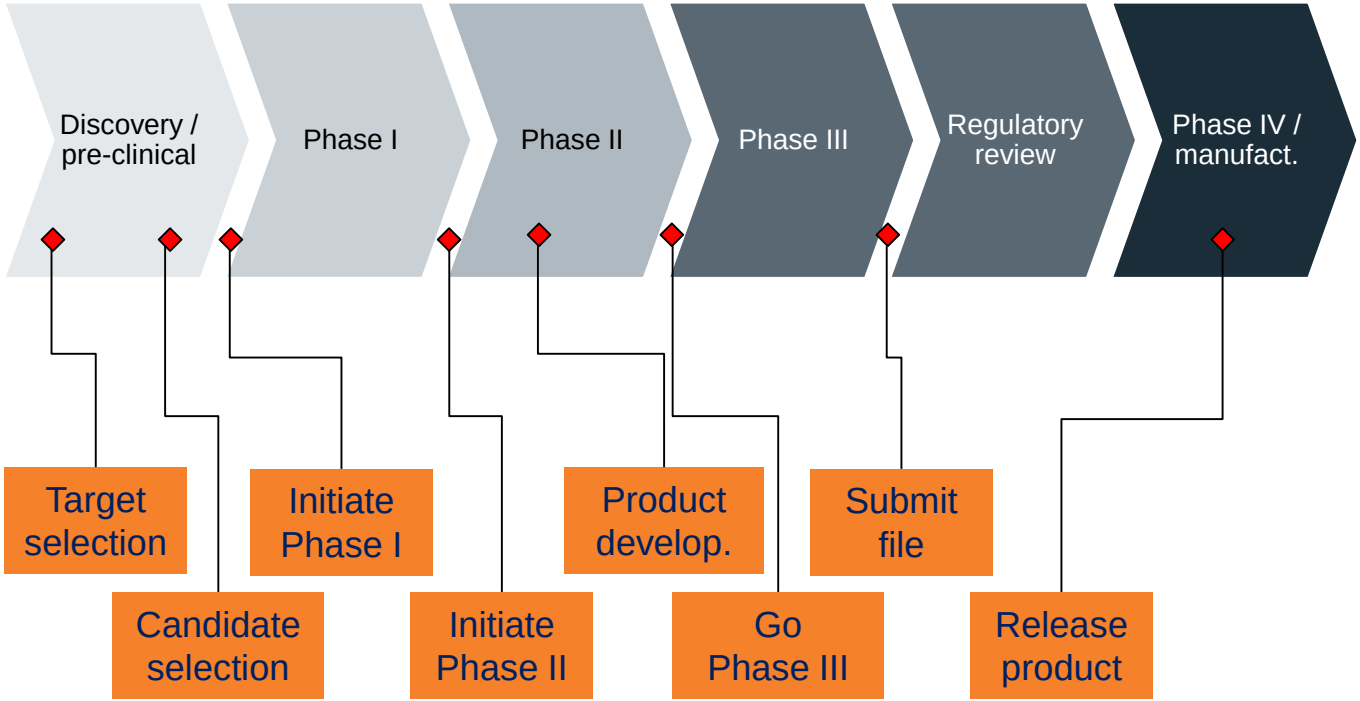

BAYES
BAYESIAN BIOSTATISTICS
LYON
2019
21-24 MAY

Drug development is a learning process



Connecting the dots:
Bayesian inference
is a learning process

Decisions through drug development and sales



The objective: is my treatment effective ?

How to make a decision ?

A

What is the probability of obtaining the observed data, if the treatment is not effective?

B

What is the probability that the treatment is effective, given the observed data?

Two different ways to make a decision based on

A

Pr(**observed data** | **treatment is not effective**)

- Better known as the **p-value** concept
- Used in the **null hypothesis** test (or decision)
- This is the likelihood of the data assuming an hypothetical explanation (e.g. the “null hypothesis”)
- **Classical statistics** perspective (Frequentist)

B

Pr(**treatment effective** | **observed data**)

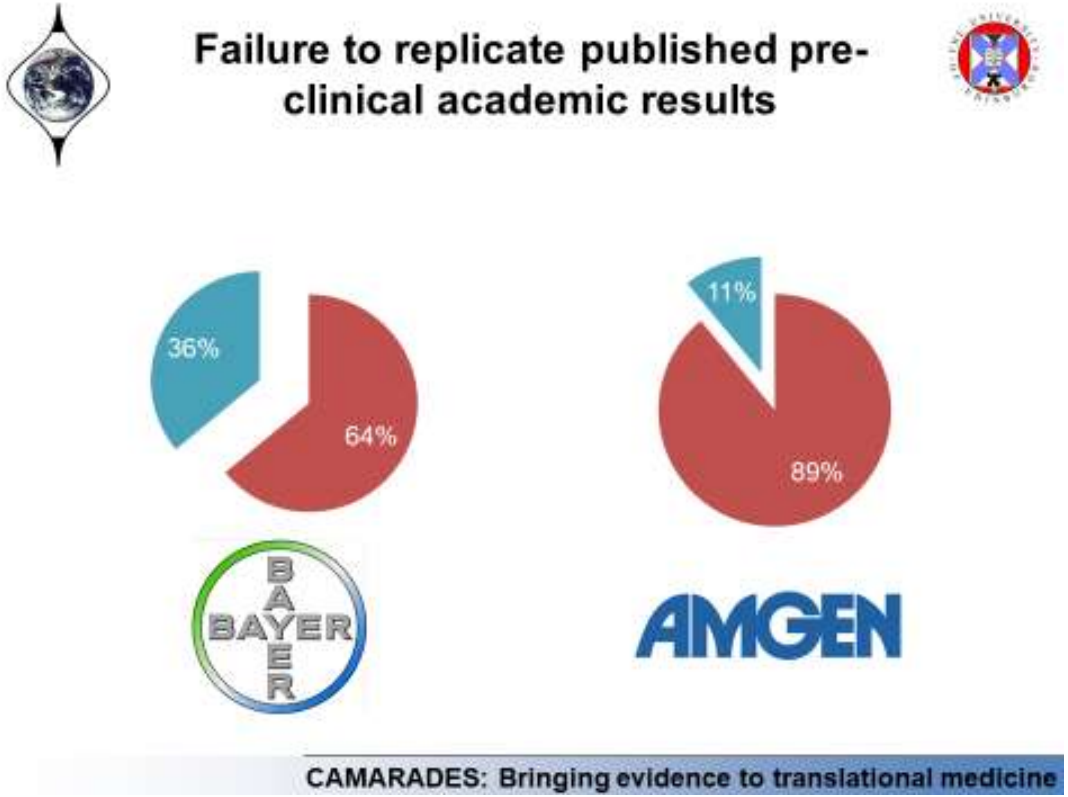
- **Bayesian** perspective
- It is the probability of efficacy given the data

The Bayesian perspective allows to directly address the question of interest.

One size does not fit all….What's the question?

Reproducibility crisis

The “Bayer” and “Amgen” publications



Nature, 2014

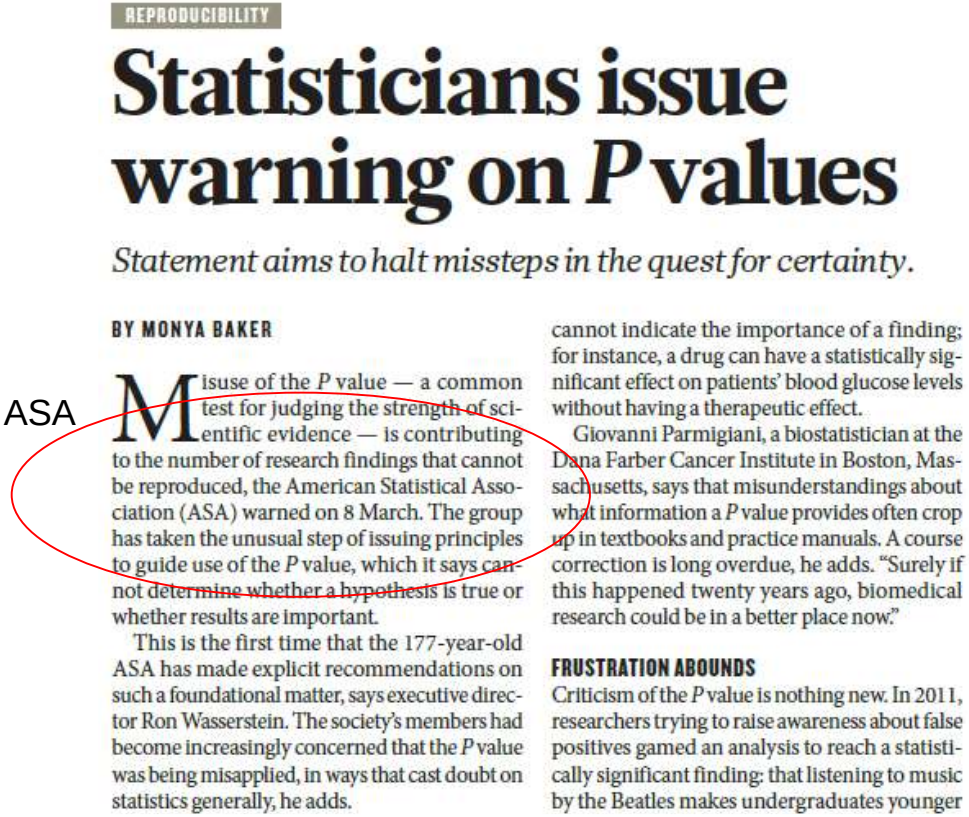
STATISTICAL ERRORS

P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume.

BY REGINA NUZZO

March 2016:

Official reminder of ASA



REPRODUCIBILITY

Statisticians issue warning on *P* values

Statement aims to halt missteps in the quest for certainty.

BY MONYA BAKER

Misuse of the *P* value — a common test for judging the strength of scientific evidence — is contributing to the number of research findings that cannot be reproduced, the American Statistical Association (ASA) warned on 8 March. The group has taken the unusual step of issuing principles to guide use of the *P* value, which it says cannot determine whether a hypothesis is true or whether results are important.

This is the first time that the 177-year-old ASA has made explicit recommendations on such a foundational matter, says executive director Ron Wasserstein. The society's members had become increasingly concerned that the *P* value was being misapplied, in ways that cast doubt on statistics generally, he adds.

cannot indicate the importance of a finding; for instance, a drug can have a statistically significant effect on patients' blood glucose levels without having a therapeutic effect.

Giovanni Parmigiani, a biostatistician at the Dana Farber Cancer Institute in Boston, Massachusetts, says that misunderstandings about what information a *P* value provides often crop up in textbooks and practice manuals. A course correction is long overdue, he adds. "Surely if this happened twenty years ago, biomedical research could be in a better place now."

FRUSTRATION ABOUNDS

Criticism of the *P* value is nothing new. In 2011, researchers trying to raise awareness about false positives gamed an analysis to reach a statistically significant finding: that listening to music by the Beatles makes undergraduates younger

Key statement from ASA Press Release

- A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
- Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
- By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

- P-values can indicate how incompatible the data are with a specified statistical model.
- P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
- Proper inference requires full reporting and transparency.

Bayesian reasoning

- The diagnostic test example

Cancer ? → diagnostic test → result

A problem of decision making

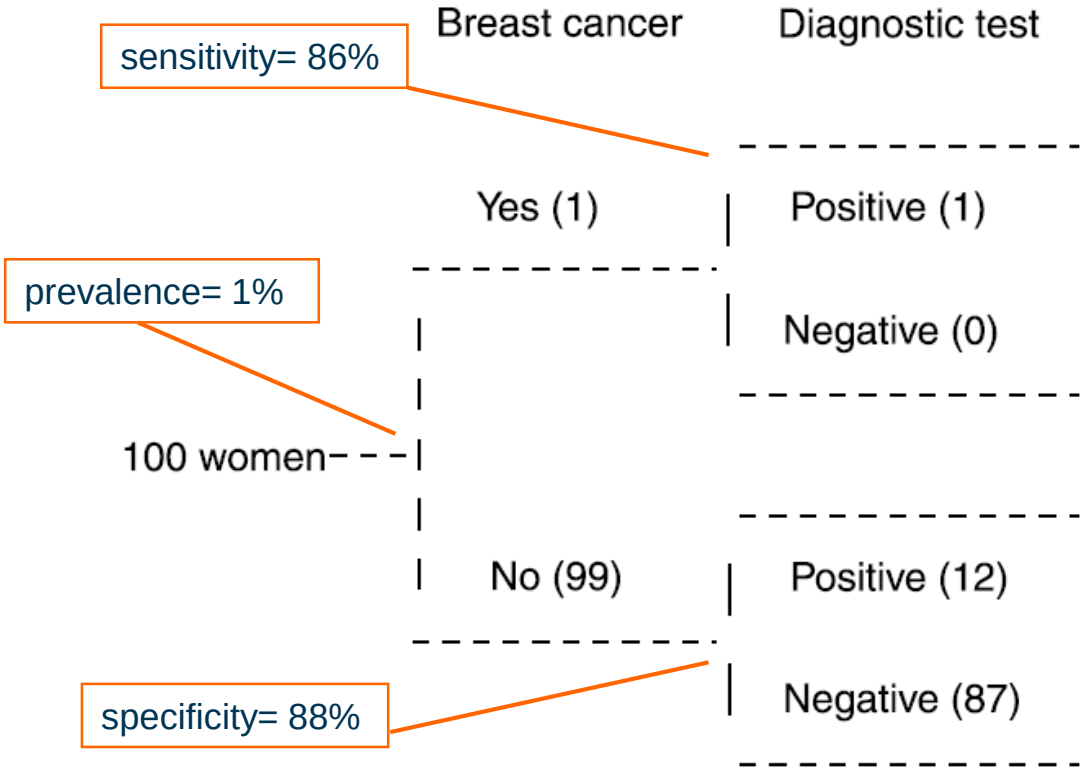
- The accuracy of a diagnostic test is assessed as follows:
  - Sensitivity: $\Pr(\textbf{positive result} \mid \textbf{cancer})$
  - Specificity: $\Pr(\textbf{negative result} \mid \textbf{no cancer})$

In practice:

Given that the diagnostic test result is positive, what is the probability you truly have cancer?

$$\Pr(\textbf{cancer} \mid \textbf{positive result}) = ?$$

Example



$$\Pr(\textbf{cancer} \mid \textbf{positive result}) = \frac{1}{12+1} = 0.077$$

How can that be so low?
The small proportion of errors for the large majority of women who do not have breast cancer swamps the large proportion of correct diagnoses for the few women who have it.

The probability of interest dependents on the underlying prevalence of the disease.

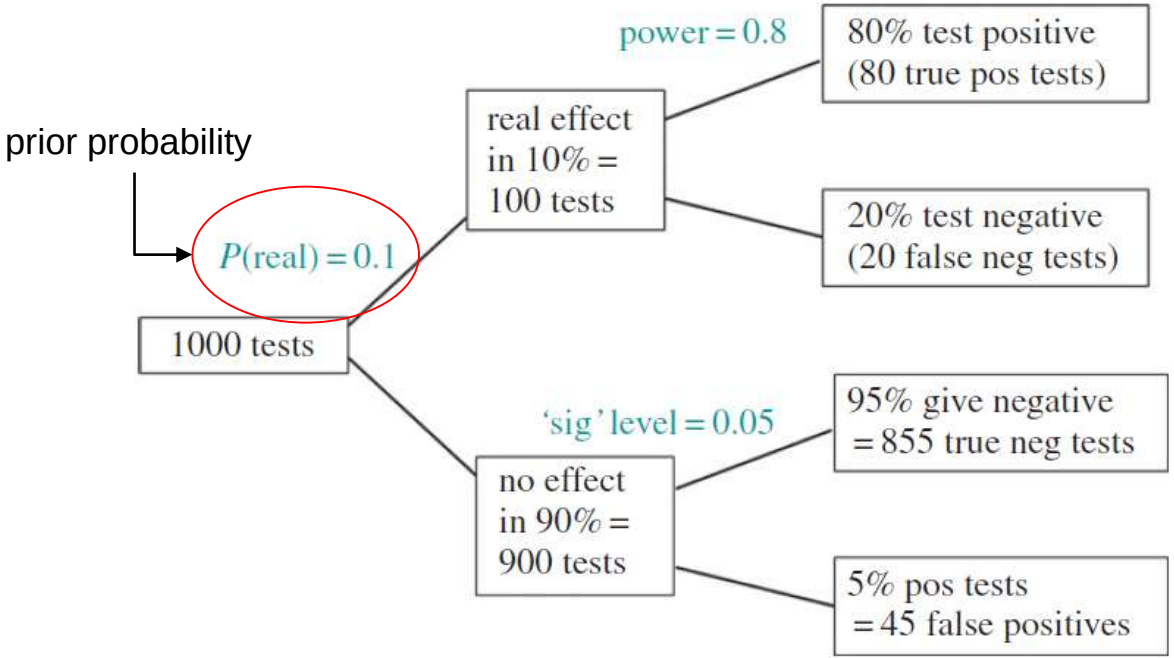
Agresti, A. (2007). *An Introduction to Categorical Data Analysis.* Wiley, 2nd ed.
Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of *p*-values. *R. Soc. Open sci.* **1**(3): 140216

The clinical trial analogy

effective? → clinical trial → data

$$\Pr(\textbf{drug effective} \mid \textbf{data}) = ?$$

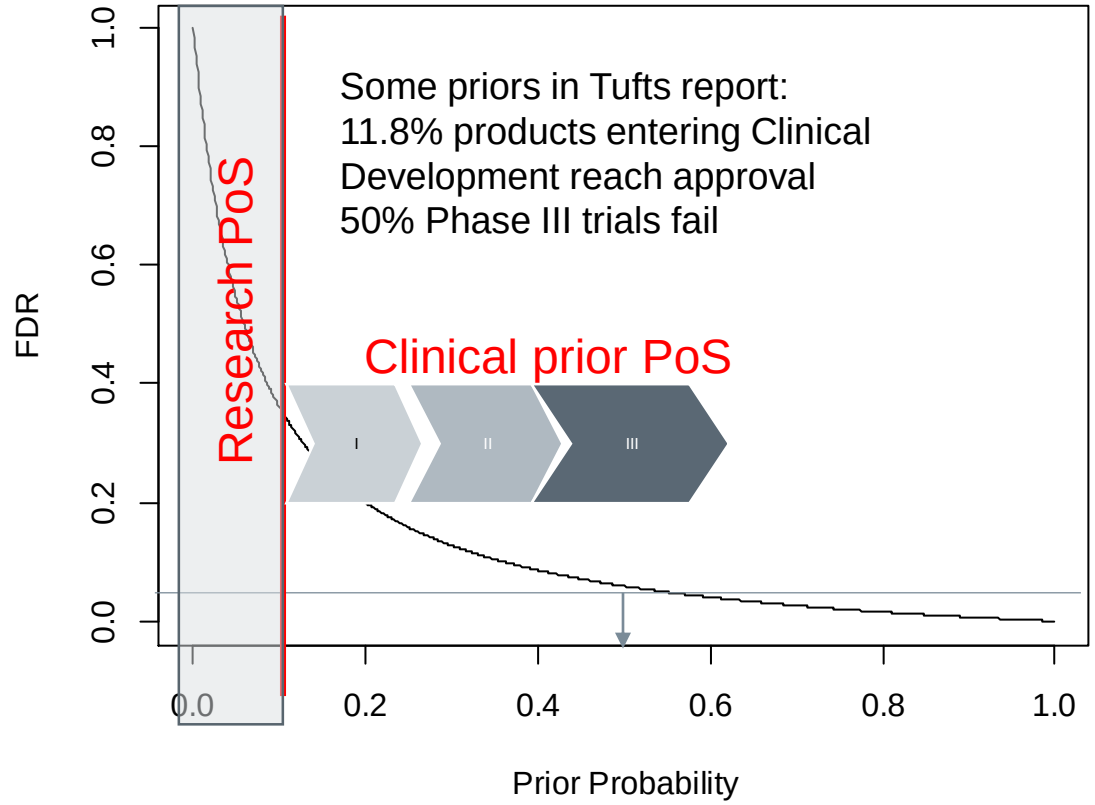
“If you use p = 0.05 to suggest that you have made a discovery, you will be wrong at least 30% of the time.”



$$\Pr(\textbf{real effect} \mid \mathbf{p} < 0.05) = \frac{80}{80 + 45} = 0.64$$

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of *p*-values. *R. Soc. Open sci.* **1**(3): 140216.

False Discovery Rate for p<0.05, power=0.8 as function of Prior Probability



Some statisticians said….

- "The most important task before us in developing statistical science is to demolish the P-value culture, which has taken root to a frightening extent in many areas of both pure and applied science and technology."
  Nelder, J. A. 1999. Statistics for the millennium. Statistician 48:257–269.

- “Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.”
  Ron Wasserstein, President American Statistical Association, March 2016

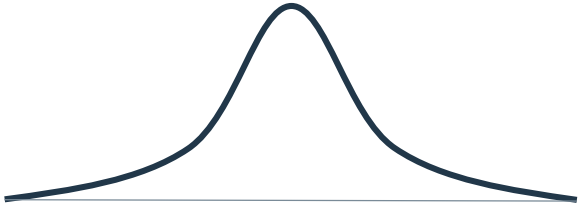
- “… we recommend abandoning the null hypothesis significance testing paradigm entirely, leaving p-values as just one of many pieces of information with no privileged role in scientific publication and decision making.”
  McShane, Gal, Gelman, Robert & Tackett, 21SEP2017

Meeting at JSM in 2016

"*FDA is facing difficulties to recruit statisticians trained in Bayesian statistics. Why do you (academic) continue to train all statisticians almost exclusively in Null Hypothesis Significance Testing ?*"

Bayesian inference is the mechanism used to update the state of knowledge

prior information
$p(\theta)$

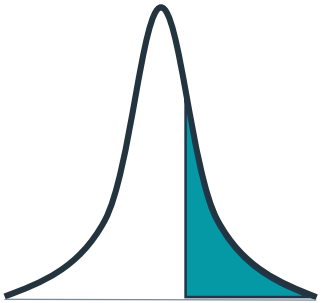


+

data information
$p(\text{data}|\theta)$

$\propto$

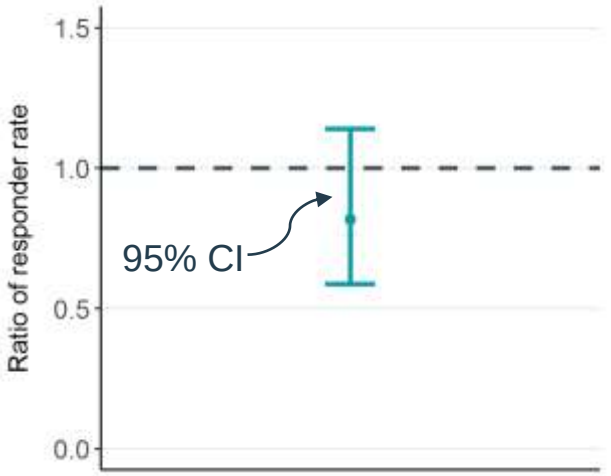
posterior information
$p(\theta|\text{data})$



The process to arrive at a posterior distribution makes use of Bayes' formula.

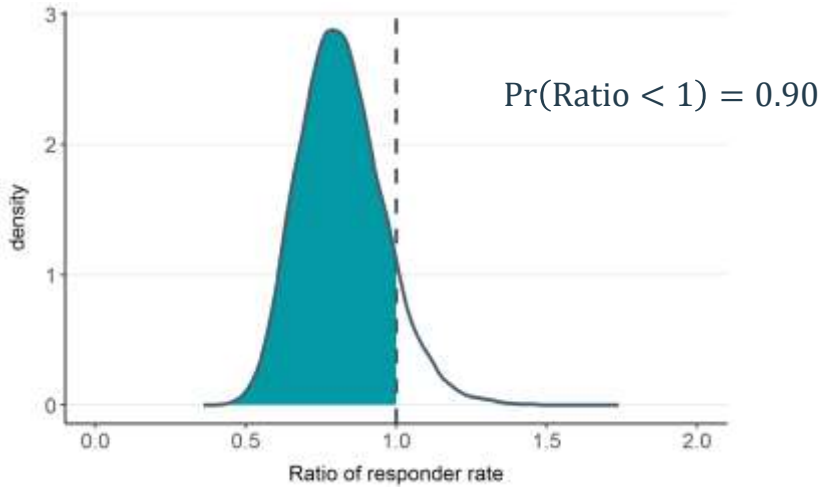
Example: Ratio of two proportions (2/2)

frequentist approach



➔ With a frequentist analysis, confidence intervals are calculated using **normal approximation**.

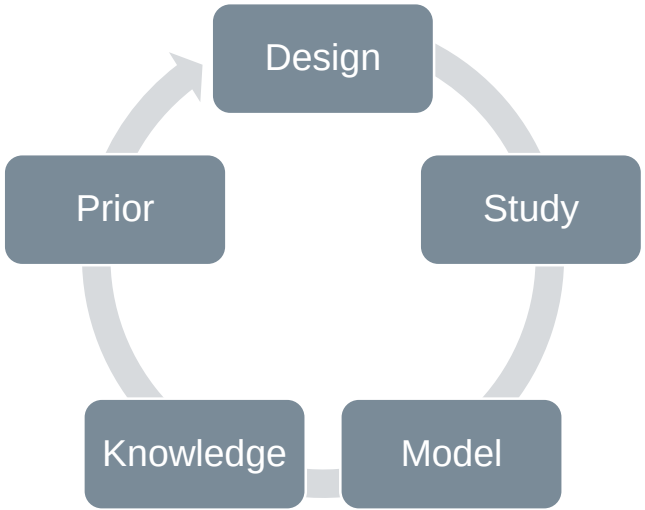
bayesian approach



➔ With a Bayesian analysis, posterior distribution of the ratio is **easily obtained from a sample of the** posterior distribution of the ratio.

Statistics-Based Drug Development with Bayesian statistics

- Bayesian statistics became popular in drug development with Adaptive Designs
- Prior knowledge is used to make the trial design more effective and informative
- It’s a key player in a **Statistical** Model-Based Drug Development strategy

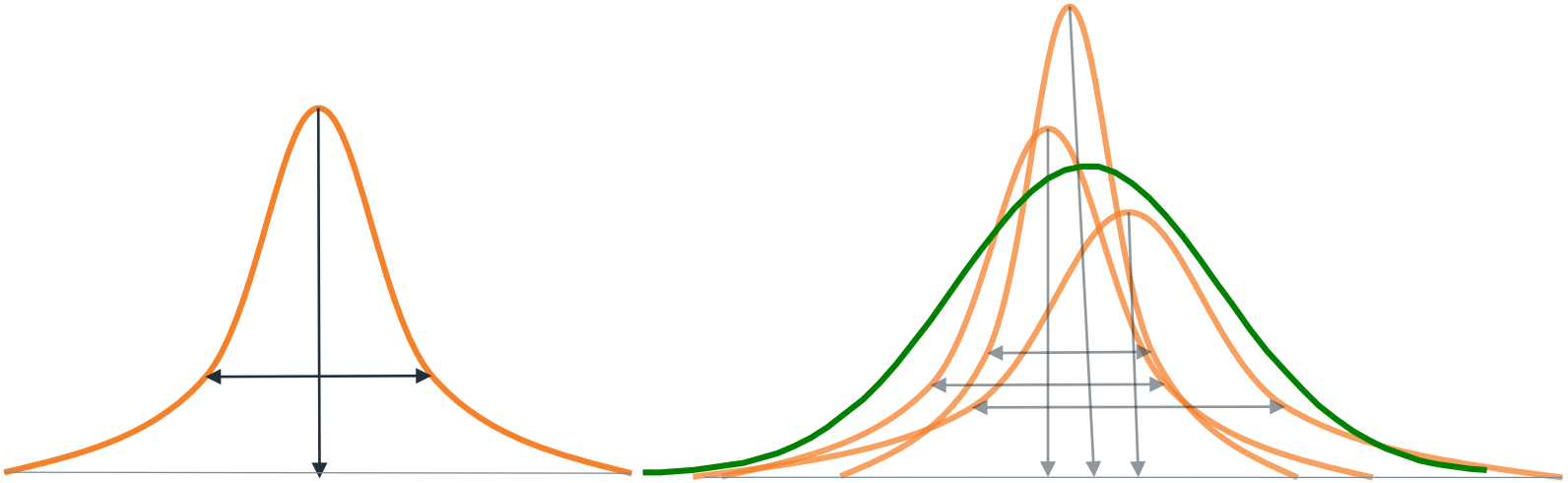


Design future trials: Simulations VS predictions

Predictions takes into account the uncertainty in estimating the model parameters.

Simulations

Predictions



Power vs assurance (1/3)

independent samples t-test ($H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$)

frequentist approach (power)

- A power calculation takes a particular value of the effect within the range of possible values given by H_1 and poses the question: if this particular value happens to obtain, what is the probability of coming to the correct conclusion that there is a difference?

assumptions:

- $\mu_1 = 100;$
- $\mu_2 = 120;$
- $\sigma_1^2 = \sigma_2^2 = 39$

very strong priors!

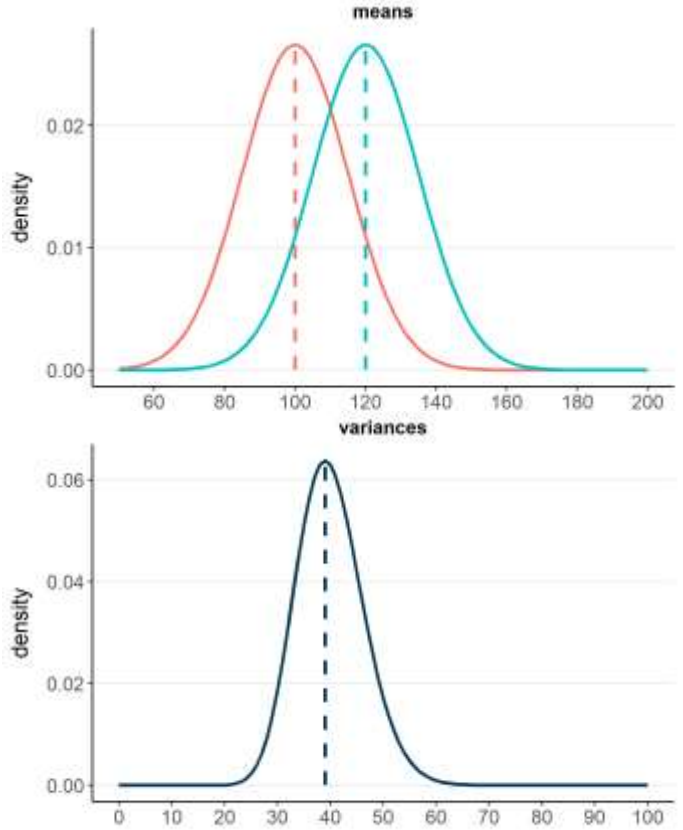
Power vs assurance (2/3)

independent samples t-test ($H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$)

bayesian approach (assurance)

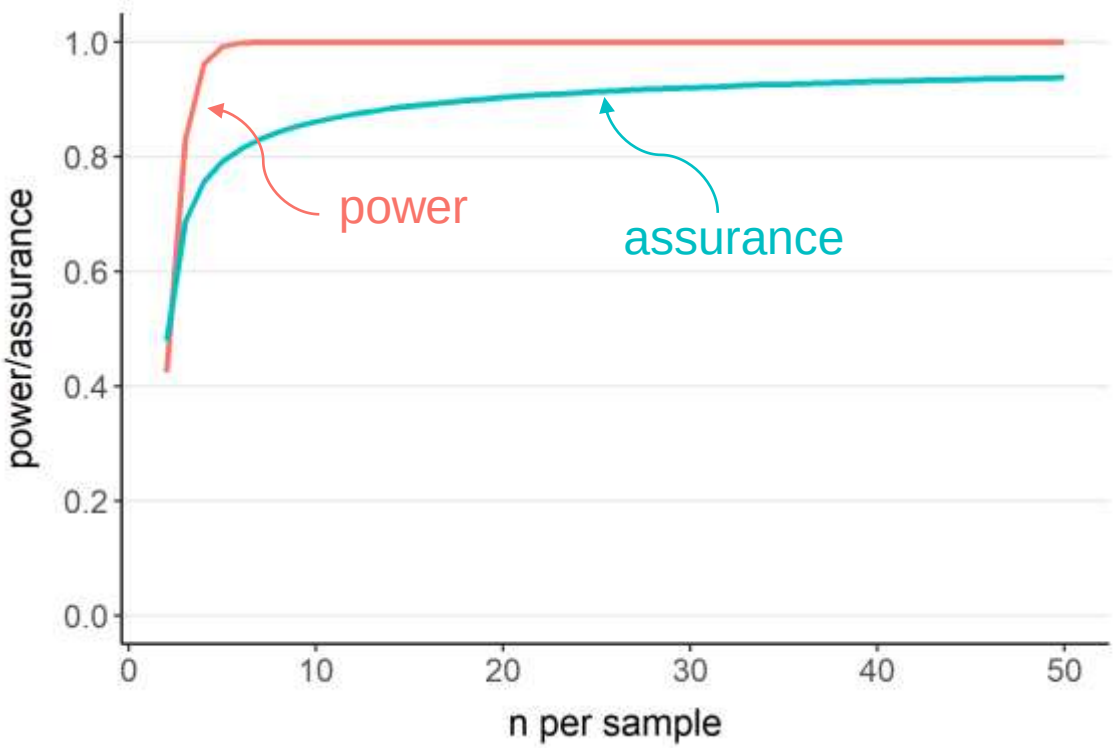
- In order to reflect the uncertainty, a large number of effect sizes, i.e. $(\mu_1 - \mu_2)/\sigma_{pooled}$, are generated using the prior distributions.
- A power curve is obtained for each effect size
- the expected (weighted by prior beliefs) power curve is calculated

assumptions:



Power vs assurance (3/3)

independent samples t-test ($H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$)

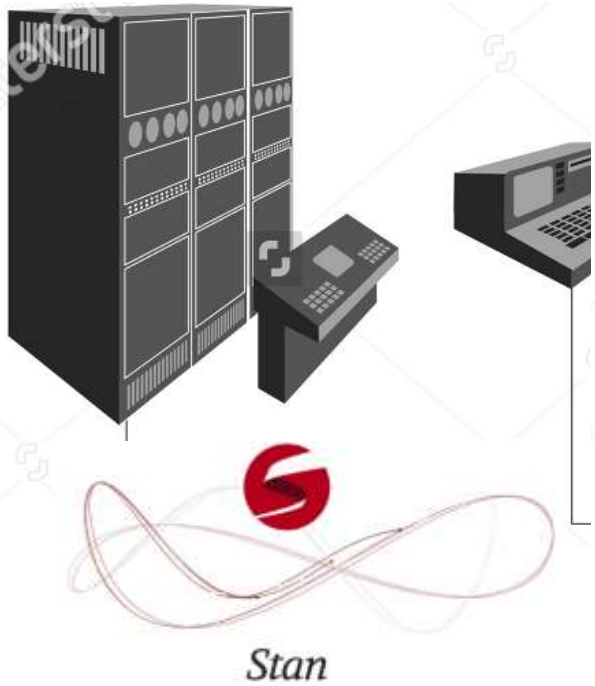


Why now and not before?

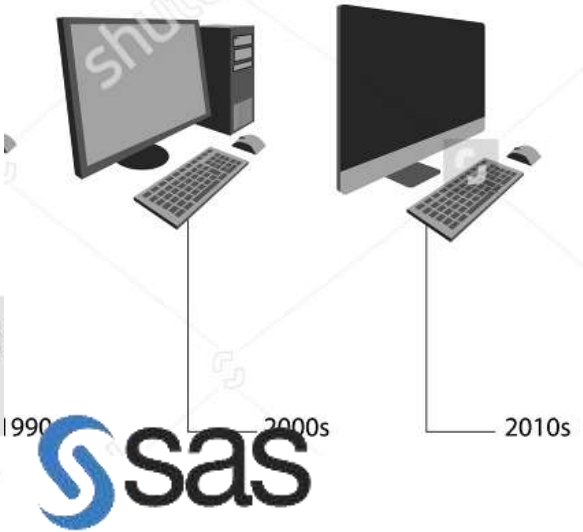
T. Bayes.

1702 - 1761

We now have computing power and more languages to apply Bayesian statistics



Stan



sas

Regulatory point of view

- 2010 - Guidance for medical device clinical trials

Guidance

for Industry and FDA Staff

Guidance for the Use of

Bayesian Statistics in

Medical Device Clinical Trials

Document issued on: February 5, 2010

2015 21st Century Cures Act US Congress Bill

A BILL

To accelerate the discovery, development, and delivery of 21st century cures, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

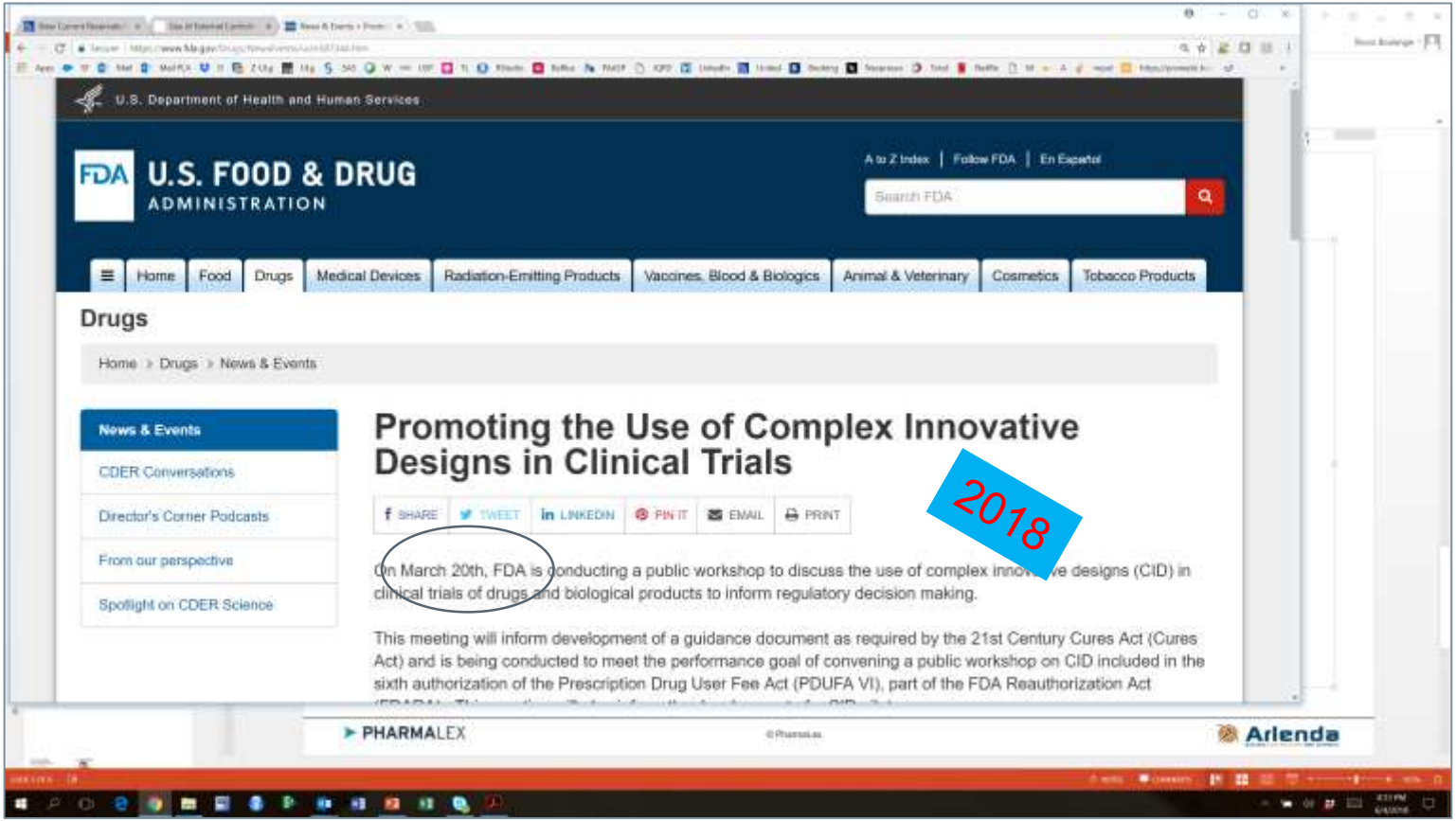
SECTION 1. SHORT TITLE.

This Act may be cited as the "21st Century Cures Act".

TITLE III—MODERNIZING CLINICAL TRIALS

Subtitle A—Clinical Research Modernization

Subtitle B—Broader Application of Bayesian Statistics and Adaptive Trial Designs


U.S. Department of Health and Human Services
FDA
U.S. FOOD & DRUG
ADMINISTRATION
A to Z Index | Follow FDA | En Español
Search FDA
Home
Food
Drugs
Medical Devices
Radiation-Emitting Products
Vaccines, Blood & Biologics
Animal & Veterinary
Cosmetics
Tobacco Products
Drugs
Home > Drugs > News & Events
News & Events
CDER Conversations
Director's Corner Podcasts
From our perspective
Spotlight on CDER Science
Promoting the Use of Complex Innovative Designs in Clinical Trials
SHARE
TWEET
LINKEDIN
PIN IT
EMAIL
PRINT
2018
On March 20th, FDA is conducting a public workshop to discuss the use of complex innovative designs (CID) in clinical trials of drugs and biological products to inform regulatory decision making.
This meeting will inform development of a guidance document as required by the 21st Century Cures Act (Cures Act) and is being conducted to meet the performance goal of convening a public workshop on CID included in the sixth authorization of the Prescription Drug User Fee Act (PDUFA VI), part of the FDA Reauthorization Act
PHARMALEX
Arlenda

U.S. FOOD & DRUG ADMINISTRATION

Complex Innovative Designs

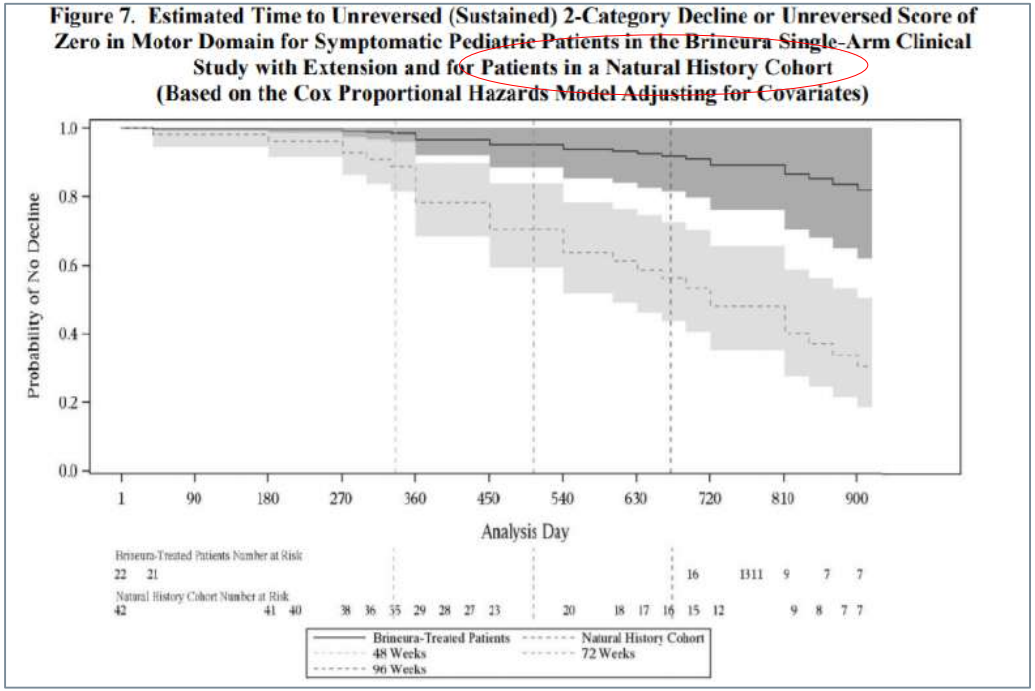
2018

- Adaptive designs that are complex, due to:
  - Adaptations on multiple factors, and/or
  - Requiring simulations to determine operating characteristics
- Other designs incorporating
  - Innovative use of external data
  - Innovative criteria for decision-making
  - Innovative collaborative efforts

Historical control

Rare Disease Example

2018



FDA discussion

U.S. FOOD & DRUG ADMINISTRATION

Bayesian Applications

2018

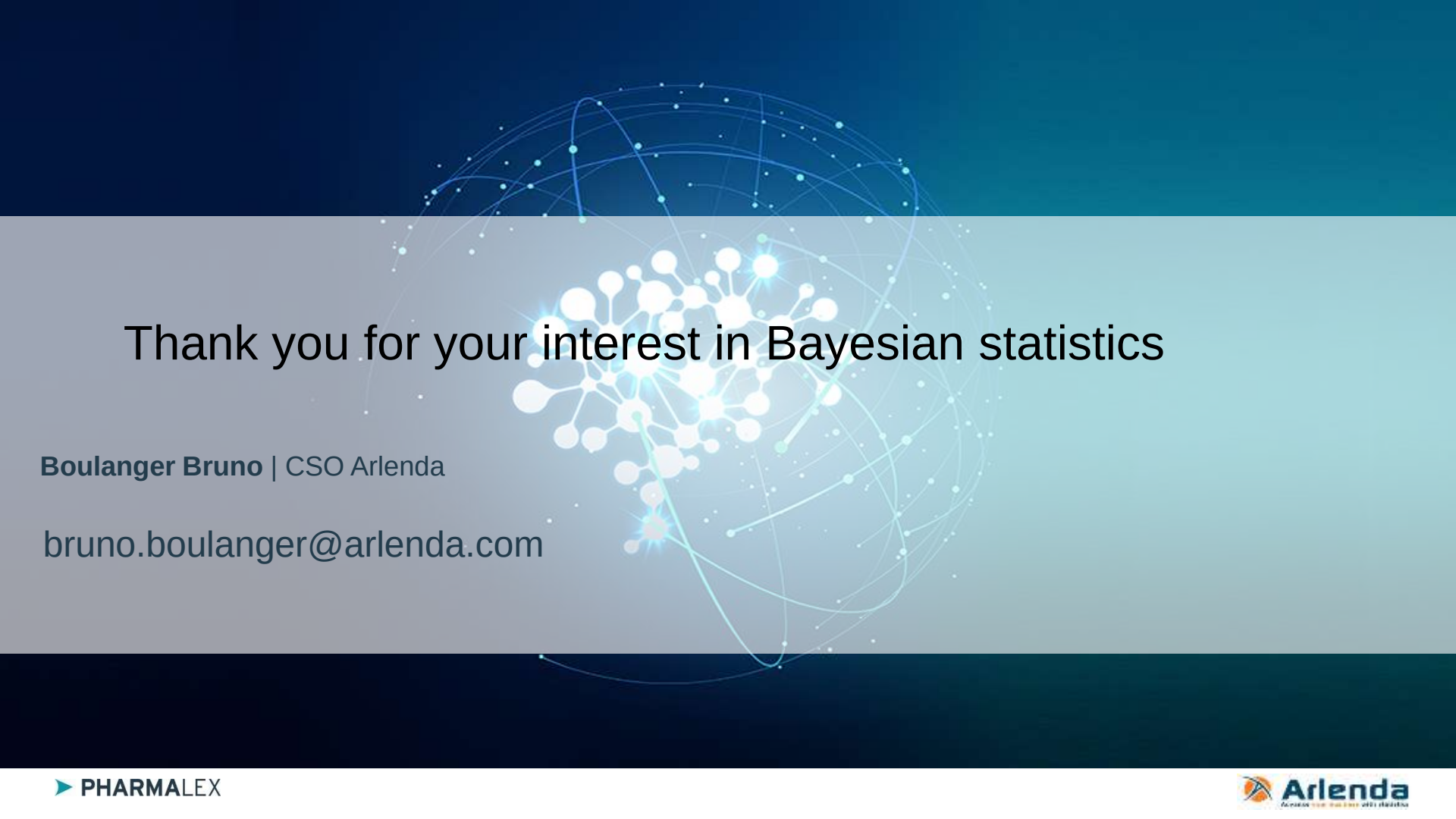
- Safety monitoring
  - Large CV risk studies that leverage control patient data from other sources via Bayesian adaptive designs
- Oncology
  - Early phase dose-finding trial designs, e.g., CRM
  - Bayesian adaptive trials that use intermediate or accelerated approval endpoints for decision-making
- Rare diseases
  - Incorporate prior information from early phase trials
  - Use information about disease progression in analytical model
  - Compute shrinkage estimators of effects in rare subsets of disease
  - Incorporate prior information from adult trials to improve efficiency of pediatric trials

Conclusions

What's Next ?

- There is a clear move for a broader use and acceptance of Bayesian statistics
- How to ensure all actors will evolve at the same pace?
  - Collaborate with Regulatory authorities seems a priority
  - Share more applied examples (vertical)
  - Extend the scope of application (horizontal)
- Actions to be taken by EFSPI
  - Already one WG ongoing
  - More focused WG ?
  - What are the priorities ?
  - How to involve EMA ?

- Your turn!



Thank you for your interest in Bayesian statistics

Boulanger Bruno | CSO Arlenda

bruno.boulanger@arlenda.com