Science For A Better Life

# Statistics for Decision Processes - Transitions between Research and Development Phases

PSI Scientific Meeting, March 30th, 2017/ Richardus Vonk, Head of Research and Clinical Sciences Statistics

# Disclaimer

The views expressed in this presentation
is the personal view of the author,
and do not necessarily represent the view of his employer

# Agenda

- Introduction
- Decision Making in PoC Studies
- MAP approach in early clinical development
- Biomarkers / Enrichment Designs
- Discussion and Conclusions

# Acknowledgments


- Heinz Delesen for work and slides on Bayesian Concepts for PoC Studies
- Rong Liu and Oliver Boix for further work on visualization for decision making
- Andreas Kaiser and Stefan Klein for work and slides on MAP approach
- Harry Mager and RCSS for discussions and continued appetite for innovation
- Parts of this project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.



# Statistics…

Statistical thinking and methods
are an integral part of the decision processes,
and form the indispensable basis
of all drug discovery and development phases

# Translational Medicine

“The threat to increased output is the fact that most activities under the umbrella "translational medicine" are pretentious and reflect phraseology, thus just wishful thinking.”

Tools of Translational Science in Medicine as backbone of an emerging science

- New biomarker development, e.g. imaging or serum parameters
  Translational toxicology including more powerful biomarkers
- Biomarker scoring systems to grade their predictive potency
- Smart, early human study design, including novel approaches e.g. microdosing and descriptive trials
- Biostatistics development to cope with multiple read-out problems and small human studies
- Human genetics

**Wehling, M (2008). Translational medicine: science or wishful thinking? Journal of Translational Medicine 6:31**

# Translational Assessment Aspects

- Starting evidence 11%
- Human evidence 13%
- Biomarkers for efficacy and safety prediction 37%
- PoM, PoP, PoC 13%
- Personalized medicine aspects 8%

Wehling (2009). Assessing the translatability of drug projects: what needs to be scored to predict success? Nat Rev Drug Discov 8:541-546.

- Biomarkers and personalized medicine aspects play an important role (~45%)
  - Biomarkers 37%
  - Biomarker strategy (PoM, PoP, PoC) 5%
  - Disease subclassification and concentration of “responders” (personalized medicine aspects 3%

# Translation from Animal Models

## Translation from Animal Models in Cancer Research[1]

- Due to practical and ethical concerns associated with human experimentation, animal models have been essential in cancer research.
- However, the average rate of successful translation from animal models to clinical cancer trials is less than 8%.

## Translatability[2]

- Suitable animal models are indispensable
- Important points to develop are: similarity to human disease, biomarker strategies bridging between animal and human studies, multiple species
- Hackam[3] (2006) “Only about a third of highly cited animal research translated at the level of human randomized trials”

1: Mak, I, Evaniew, N, Ghert, M (2014). Lost in translation: animal models and clinical trials in cancer treatment. Am J Trans Res 6: 114-118

2: Wendler, A, Wehling, M (2010). The translatability of animal models for clinical development: biomarkers and disease models. Current opinions in pharmacology 10:601-606.

3: Hackam DG, Redelmeier DA (2006) Translation of research evidence from animals to humans. JAMA 296: 1731–1732.

# Statistics in Translation
## From Mice To Men

Evaluation of 271 articles, live rats, mice and non-human primates carried out in UK and US publicly funded research establishments[1]

- "59% stated hypothesis or objectives of the study and the number and characteristics of the animals used"
- Randomization: 13%, blinding: 14%
- 70% of the papers that used statistical methods described these, and presented the results with a measure of error.

Toxicological transition into humans[2]

- 71% true positive human toxicity concordance for rodents / non-rodents
- High concordance in hematological, gastrointestinal, cardiovascular
- Low concordance in cutaneous human toxicity


1: Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, et al. (2009) Survey of the Quality of Experimental Design, Statistical Analysis and Reporting of Research Using Animals. PLoS ONE 4(11): e7824. doi:10.1371/journal.pone.0007824

2: Olson, H, et al. (2000). Concordance of the toxicity of pharmaceuticals in humans and in animals. Regul Toxicol Pharmacol 32: 56-67

# Believe it or not….

CORRESPONDENCE

LINK TO ORIGINAL ARTICLE

## Believe it or not: how much can we rely on published data on potential drug targets?

*Florian Prinz, Thomas Schlange and Khusru Asadullah*

results that are published are hard to reproduce. However, there is an imbalance between this apparently widespread impression and its public recognition (for example, see REFS 2,3), and the surprisingly few scientific publications dealing with this topic. Indeed, to our knowledge, so far there has been no published in-depth, systematic analysis that compares reproduced results with published results for wet-lab experiments related to target identification and validation.

- ~20 – 25% of projects where in-house data were completely in line with published data
- Almost 2/3 of the projects showed inconsistencies which considerably prolonged the duration of target validation or, in most cases, resulted in termination of the project

Prinz, F., Schlange, T., Asadullah, K. (2011). Believe it or not: How much can we rely on published data on potential drug targets? *Nature Review Drug Discovery* 10:712–713.

# Raise the Standards?

Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

## Raise standards for preclinical cancer research

**C. Glenn Begley** and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.

Efforts over the past decade to characterize the genetic alterations | trials in oncology have the highest failure rate compared with other therapeutic areas | investigators must reassess their approach to translating discovery research into greater

Amgen in 2012

Scientific findings were confirmed in only 6 out of 53 papers (11%)

Begley CG, Ellis LM (2012) Drug development: raise standards for preclinical cancer research. Nature 483: 531–533. doi: 10.1038/483531a PMID: 22460880

# Reproducibility of Research – Consolidation in 2015

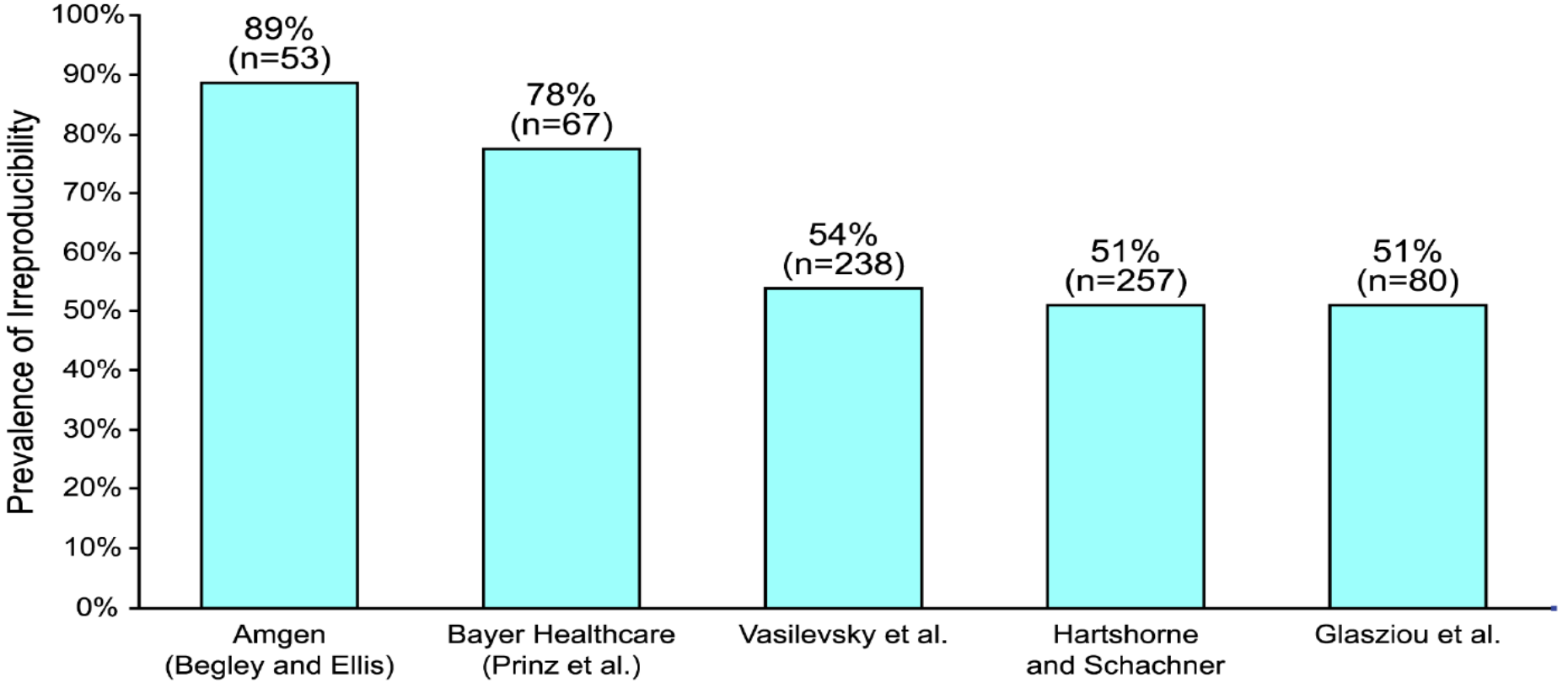



**Fig 1. Studies reporting the prevalence of irreproducibility.** Source: Begley and Ellis [6], Prinz et al. [7], Vasilevsky [8], Hartshorne and Schachner [5], and Glasziou et al. [9].

Freedman LP, Cockburn IM, Simcoe TS (2015) The Economics of Reproducibility in Preclinical Research. PLoS Biol 13(6): e1002165.doi:10.1371/journal.pbio.1002165

# Back to Translation towards Clinical Development

## General

- Harmonization of experimental settings between clinical and research experiments
  - Ensure that the measurements in research are aligned with those in clinical development
- Ensure knowledge transfer – not only about the compound, but also about experimental setting – in both directions

## Challenges

- Harmonization of (statistical) methodology used in research and clinical development
- Determination of scalability of results from research to clinical application
- Deal with differences between species
- Communication between pre-clinical research and clinical development

# Failed Animal to Human Translation

Additional reasons for failed translations might be:

- Lack of internal validity of animal studies, due to, e.g., randomization, blinding, sample size calculation, eligibility criteria, statistical analysis, etc.
- Reduced external validity of animal studies, due to, e.g., differences in outcome measurements, homogenous animals vs. heterogeneous patients, single gender analysis in animals studies for disease that might occur in males and females,
- Unrealistic settings of animal studies wrt time, doses, etc.
- Shortcomings of clinical trials: Failing to acknowledge and to transport the characteristics and limitations of animal models, etc.

Van der Worp, H, Howells, D, Sena, E, Porritt, M, Rewell, S, O'Collins, V, Macleod, M. (2010): Can animal models of disease reliably inform human studies. Plos Med 7: e1000245.

# Statistical Reasoning
## In Early Clinical Development

Moving towards quantitative decision making for PoC

- Quantitative techniques help to consider different scenarios earlier in the project
  - Earlier accumulation of quantitative knowledge
  - Clearer risk / benefit evaluation
  - Increased level of confidence
  - Guides translational efforts between preclinical and clinical phases of drug development
- More focus on estimation of effect sizes and variability in addition to statistical testing
- Requires up-to-date statistical techniques
  - Bayesian methodology
  - Sequential / Adaptive designs
  - Simulation techniques

# Opportunities

Quantitative rather than qualitative Research and Development

- Increased use of estimates and specification of (un-)certainty allows better planning for future trials in early and late stage development
- Increased use of Bayesian methods to quantify “risks and opportunities” for PoC decisions and beyond
- More use of available data: simulations, more exploratory analyses, meta-analytical approaches

# Bayes…

Bayes:

- General: P(Hypothesis | Data) instead of P(Data | Hypothesis)
- Requires elicitation of prior information
  - Sensitivity against choice of priors
  - Do you dare to use informative priors?
- Use in early clinical development
  - CRM designs
  - Decision for PoC
  - …

# Proof of Concept Studies
## And Bayes

Proof of concept (PoC) studies are generally dealing with one-sided hypotheses. Without loss of generality ('symmetry'), hypotheses of the form $H_0$: $\theta \leq \theta_0$ and $H_1$: $\theta > \theta_0$ will be considered in the following.

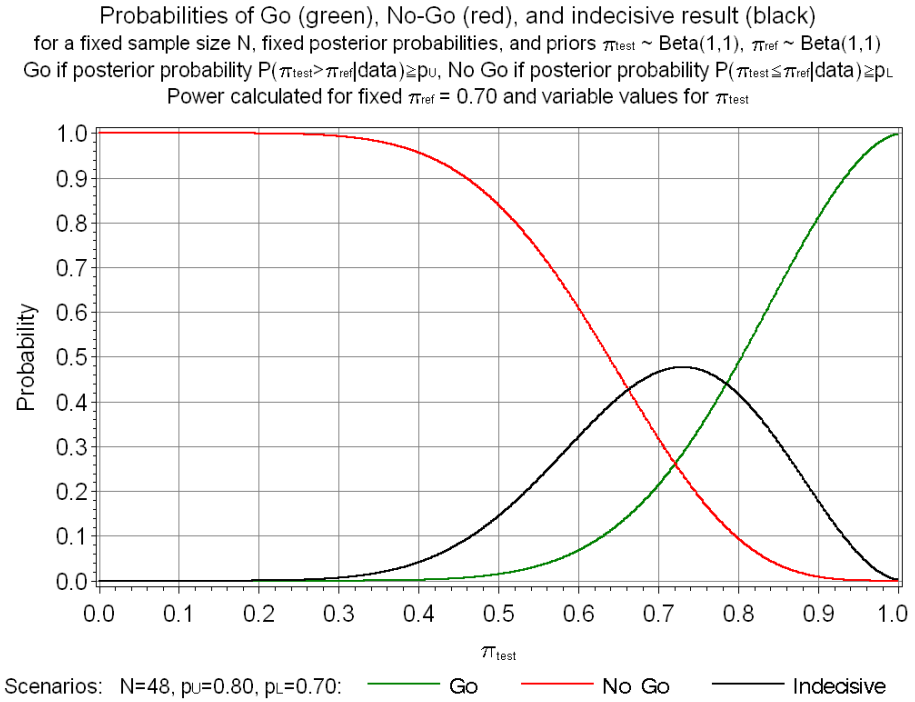
The general idea is

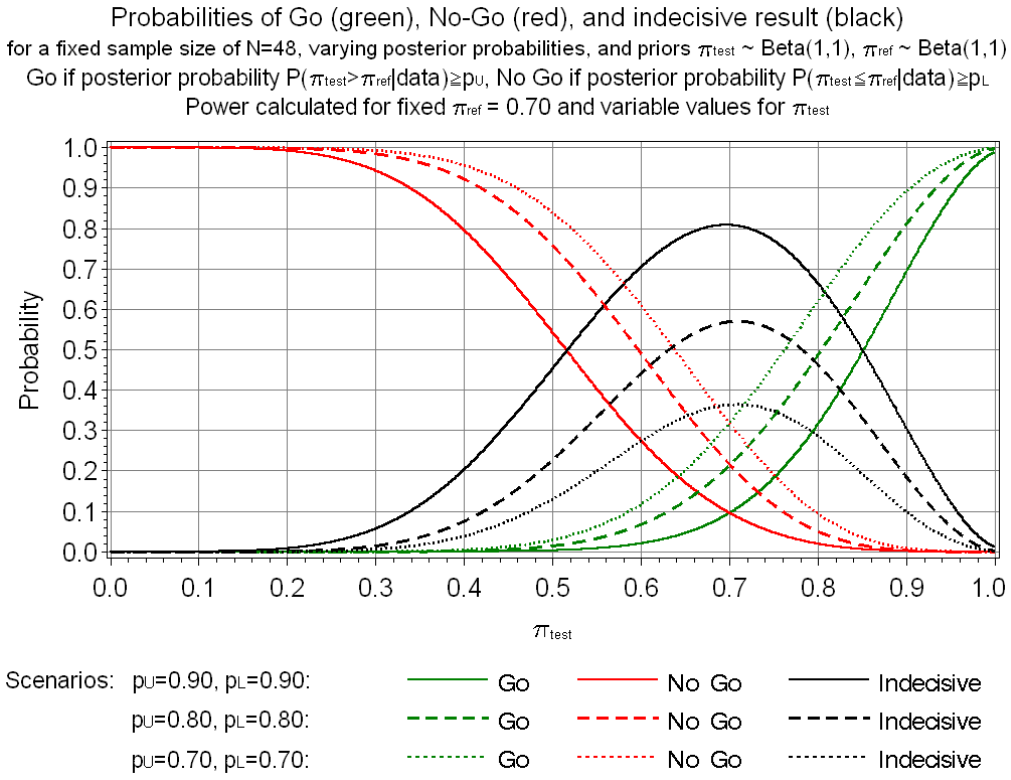
- to have a 'Go' decision if the posterior probability of $\theta > \theta_0$ is greater or equal than some pre-specified probability $p_U$,
- to have a 'No Go' decision if the posterior probability of $\theta \leq \theta_0$ is greater or equal than some pre-specified probability $p_L$,
- to have an 'indecisive' result if none of the two posterior probabilities is high enough.

$$P(\theta > \theta_0 \mid data) \begin{cases} \geq p_U \rightarrow H_1 \ (\text{"}Go\text{"}) \\ \leq 1 - p_L \rightarrow H_0 \ (\text{"}No\ Go\text{"}) \\ else \rightarrow \text{'}indecisive\text{'} \end{cases}$$

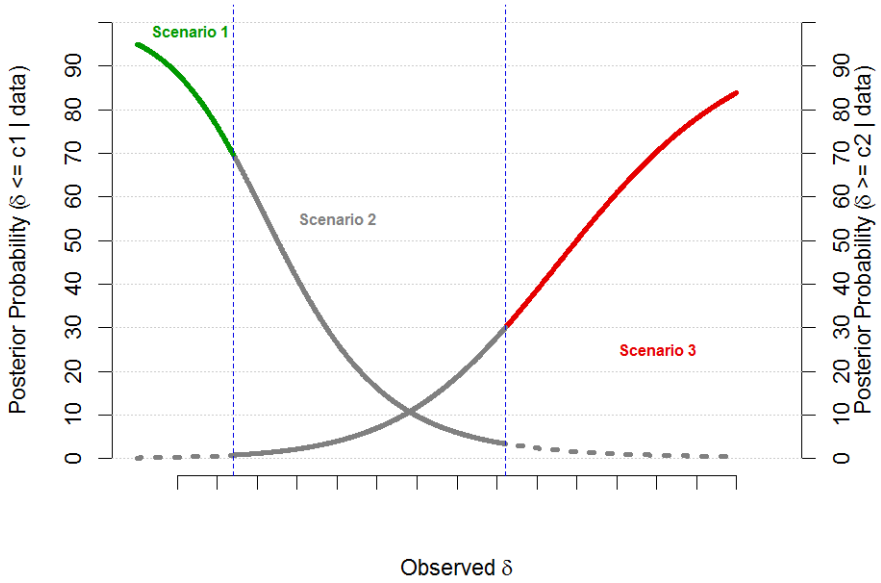
# PoC Design Properties
## Visualization

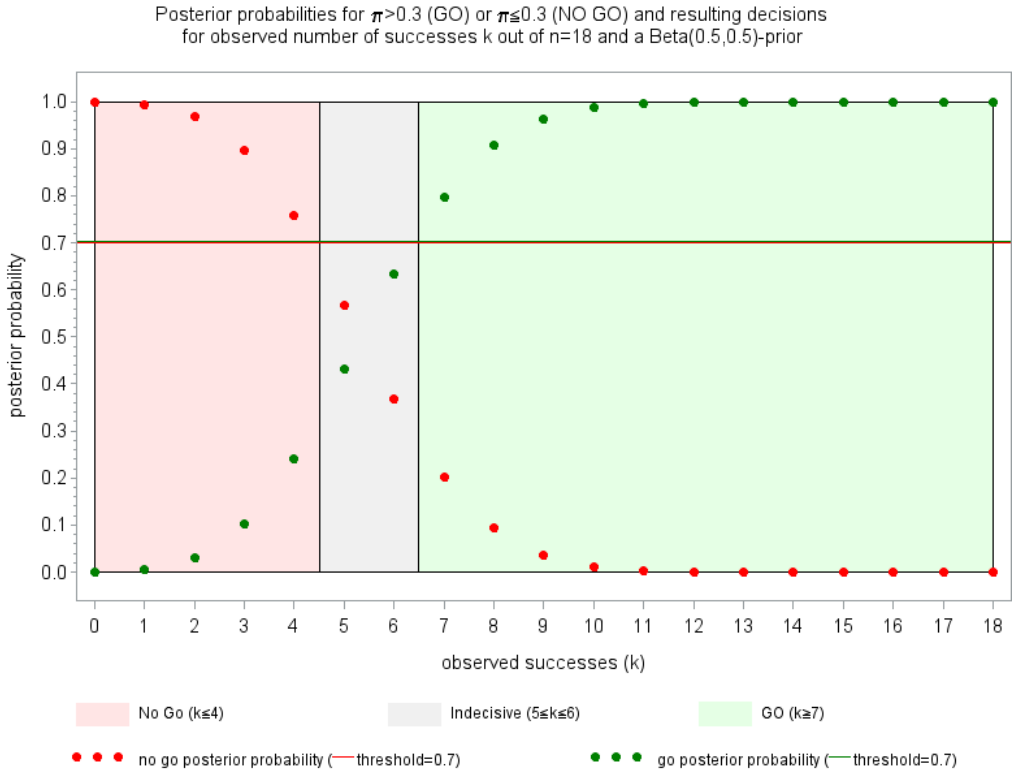
- Standard display of design properties



# Decision Making
## Visualization



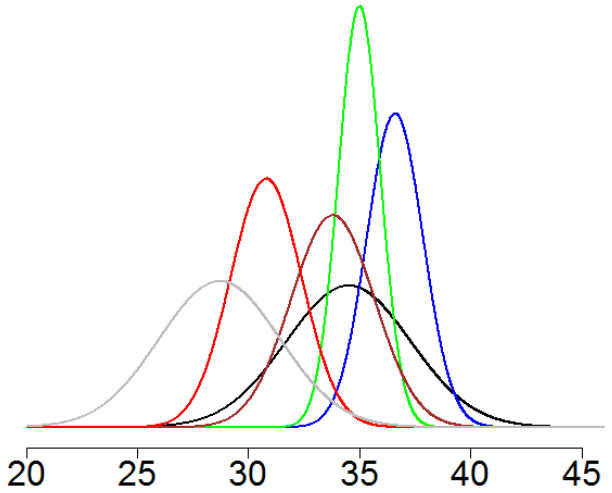
# Meta-Analytic Predictive Approach
## Application

- Introduced formally by Neuenschwander et al. (2010), but similar methods were described already in Spiegelhalter et al. (2004)
- General idea
  - Starting point: mean and SD of historical studies
  - Variability of historical studies to be decomposed into two sources: **between-trial** and **within-trial** variability
  - **Between trial variability**: nuisance parameter, but to be taken into account
  - Perform a random effects meta analysis  to assess sources of variability
  - Determine the predictive distribution for a new study and use it as a prior distribution
- Application
  - Currently applied routinely in several endpoints to assess prior distribution for (placebo or active) control arms using R programs
  - Usage of Bayesian meta analytic approaches as well as ‚normal' random effects meta analysis
  - Main outcome parameter: Effective sample size

# MAP: Dose Finding

- Study Design:
  - Phase IIb dose finding study:
    4 doses vs. active control, each 30 patients
  - primary variable: approx. normally distributed)

- Prior Information
  - 6 studies with sample sizes between 28 and 471 patients (overall: N=974)
  - Effective sample size: 80 subjects
  - Prior distribution for active control: normal distribution with $\mu$=35 and $\sigma$=20, weighted as coming from 45 patients




- Outcome
  - Smaller than maximum ESS used in order to get substantial influence from actual study data.
  - (Mean) Power increase of 10%
  - FDA: “The proposed Bayesian statistical approach … is acceptable”

# Informative Priors
## Advantages and Challenges


- Advantages
  - Saving patients by up to 30% (depending of amount of incorporated information)
  - Increase of power for decision making by up to 10%
  - Higher precision in estimation or treatment effects and model parameters
  - Increased numerical stability when estimating complex models
  - Better assessment of current trial outcome in context of historical trials
  - Better overview and more scientific discussion about realistic scenarios for trial planning
  - Positive experience regarding interaction with health authorities
- Challenges
  - Systematic deviation between study data (measurement methods, assays, endpoint definitions, population, in- and exclusion criteria, disease categories, standard of care, …)
  - Between-trial variability
  - Selection bias
  - Amount of literature available for prior derivation
  - Derivation of prior information for model parameters from published response data

# MAP and Informative Priors

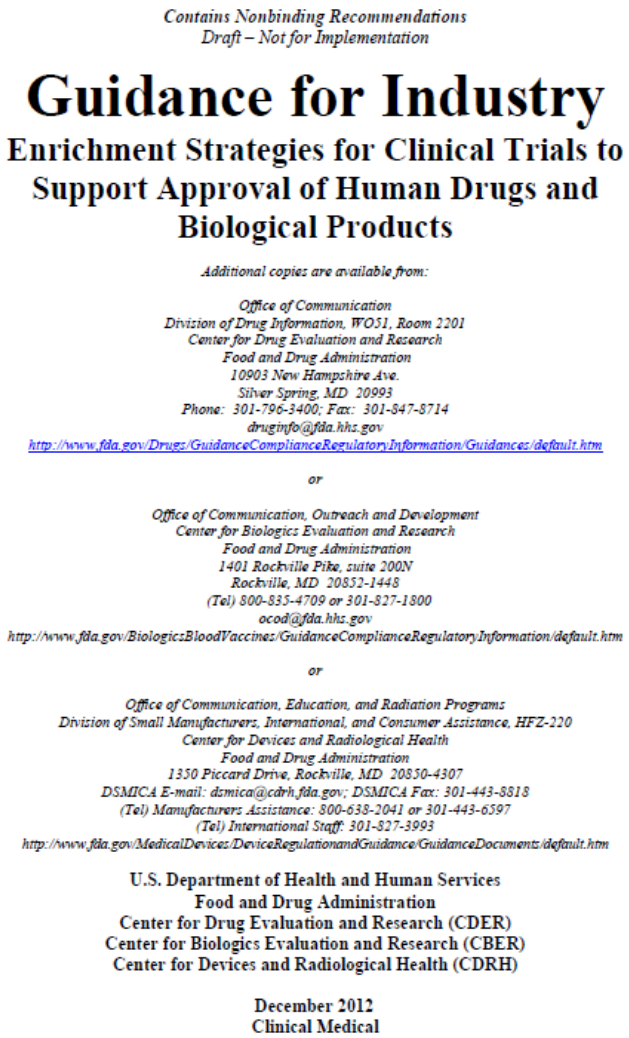
Pooling of historical data

- Down weighting necessary to cope with between-trial variability
  - Enlarging the variability of prior distribution / power priors
  - Challenge: unknown parameter for down weighting
  - Robust priors (Challenge: unknown weight for mixing distribution)

Meta-Analytical Prediction

- Able to cope with between-trial variability
- Leads to a more agreeable prior
- Challenge: Low amount of extracted information, effective sample size often ≤ 10% of overall N
- Challenge: Improvement in information extraction possible?

# Biomarkers / Enrichment Strategies



*Contains Nonbinding Recommendations*
*Draft – Not for Implementation*

**Guidance for Industry**

**Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products**

*Additional copies are available from:*

*Office of Communication*
*Division of Drug Information, WO51, Room 2201*
*Center for Drug Evaluation and Research*
*Food and Drug Administration*
*10903 New Hampshire Ave.*
*Silver Spring, MD 20993*
*Phone: 301-796-3400; Fax: 301-847-8714*
*druginfo@fda.hhs.gov*
*http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm*

*or*

*Office of Communication, Outreach and Development*
*Center for Biologics Evaluation and Research*
*Food and Drug Administration*
*1401 Rockville Pike, suite 200N*
*Rockville, MD 20852-1448*
*(Tel) 800-835-4709 or 301-827-1800*
*ocod@fda.hhs.gov*
*http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/default.htm*

*or*

*Office of Communication, Education, and Radiation Programs*
*Division of Small Manufacturers, International, and Consumer Assistance, HFZ-220*
*Center for Devices and Radiological Health*
*Food and Drug Administration*
*1350 Piccard Drive, Rockville, MD 20850-4307*
*DSMICA E-mail: dsmica@cdrh.fda.gov; DSMICA Fax: 301-443-8818*
*(Tel) Manufacturers Assistance: 800-638-2041 or 301-443-6597*
*(Tel) International Staff: 301-827-3993*
*http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/default.htm*

**U.S. Department of Health and Human Services**
**Food and Drug Administration**
**Center for Drug Evaluation and Research (CDER)**
**Center for Biologics Evaluation and Research (CBER)**
**Center for Devices and Radiological Health (CDRH)**

**December 2012**
**Clinical Medical**

Early development of biomarkers

- Find the answers to the questions, not the questions to the answers
  - Account for false positive findings
  - Validation!
- Consider loop back in translational process
- When using biomarkers for patient selections (enrichment), account for variability
  - Large(r) likely ranges of screening numbers and recruitment time if used for early enrichment

# Pharmacogenomics

## Guidance for Industry

### Clinical Pharmacogenomics: Premarket Evaluation in Early-Phase Clinical Studies and Recommendations for Labeling

*Additional copies are available from:*
*Office of Communications*
*Division of Drug Information, WO51, Room 2201*
*10903 New Hampshire Ave.*
*Silver Spring, MD 20993-0002*
*Phone: 301-796-3400; Fax 301-847-8714*
*http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm*
*or*
*Office of Communication, Outreach, and Development (HFM-40)*
*Center for Biologics Evaluation and Research*
*Food and Drug Administration*
*1401 Rockville Pike, Rockville, MD 20852-1448*
*http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm*
*(Tel) 800-835-4709 or 301-827-1800*
*or*
*Office of Communication, Education, and Radiation Programs*
*Division of Small Manufacturers, International and Consumer Assistance*
*Center for Devices and Radiological Health*
*Food and Drug Administration*
*10903 New Hampshire Ave.*
*WO66, Room 4613*
*Silver Spring, MD 20993-0002*
*http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/default.htm*
*Email: dsmica@cdrh.fda.gov*
*Fax: 301-827-8149*
*(Tel) Manufacturers Assistance: 800-638-2041 or 301-796-7100*
*(Tel) International Staff: 301-796-5708*

**U.S. Department of Health and Human Services**
**Food and Drug Administration**
**Center for Drug Evaluation and Research (CDER)**
**Center for Biologics Evaluation and Research (CBER)**
**Center for Devices and Radiological Health (CDRH)**

**January 2013**
**Clinical Pharmacology**
**Clinical/Medical**

"Statistical considerations in PGx studies are important. The hypotheses and conclusions arising from early-phase studies… should be sufficiently supported with credible data."

# Decision Making

- “There is a strong need for biologically relevant, powerful computational methods and models to integrate multi-level genome-wide evidence and to interpret the resulting high-dimensional outcomes so that strategies for clinical implementation can be developed. The major fundamental statistical challenges occur at the data analytic stage, where diverse data elements from all sources need to be incorporated into comprehensive models for prediction, risk assessment, and/or efficiency.”
- Paradigm change?
  - Move to continuous integration of data across studies
  - Much more quantitative rather than qualitative assessments, also for biomarkers
  - Highly interdisciplinary

# From Qualitative to Quantitative

From Qualitative to Quantitative Decision Making for Biomarkers

- Find Answers to the Questions, rather than Creating Questions to the Answer
  - Appropriately account for false positive findings
  - Validation!
- Consider loop back in translational process
- Use of Bayesian Methods for early discovery process

# Targeted Clinical Trials

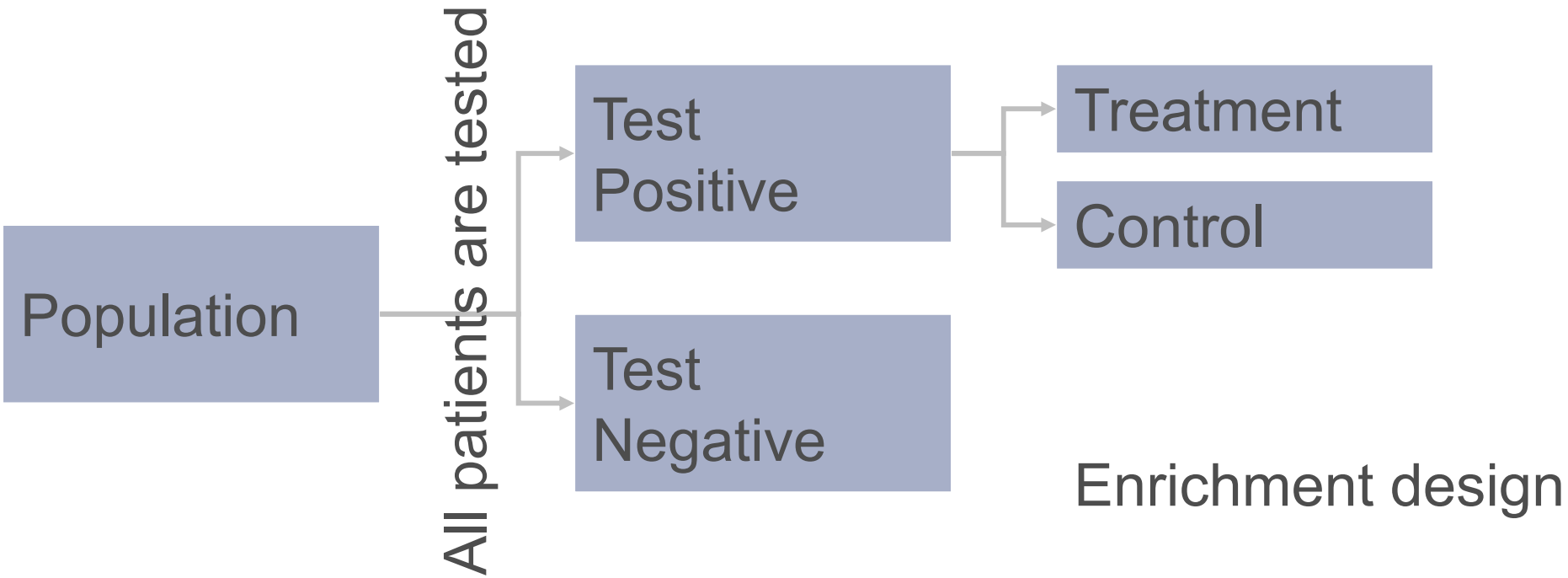
Targeted clinical trials:

- Evaluation of efficacy and safety for patient with certain (biomarker) characteristics – “biomarker positive patients”
- Evaluation of best treatment regimen depending on prognostics of clinical outcome
- Investigation of association of treatment effect with test results

# Enrichment Design

Targeted clinical trials:

- Evaluation of efficacy and safety for patient with certain (biomarker) characteristics – “biomarker positive patients”



# Challenges

High degree of certainty that relevant drug response only occurs in “marker positive” patients

- Exclusion of test negative patients prevents description of test characteristics (sensitivity, specificity)
- Effects on drug development:
  - Lower number of patients necessary,
  - but potentially longer recruitment times,
  - and potentially higher costs (screening)

Further, more generic, challenges:

- Estimation of recruitment rates
- Estimation of prevalence
- …

# Recruitment Times

- Incorporation of uncertainty increases the likely range of recruitment times – to the positive and to the negative.

- Example: recruitment rate of 10 / month, prevalence 0.5, $n^+ = 50$
- All fixed: Estimated n: 100, 90%PI [84, 117],
  Estimated time: 10 months 90%PI [7.79, 12.43]

# Effect on Screening

- Accounting for variability in estimation of prevalence:
  (note: variability in recruitment rate does not have an effect here)

| Prevalence | Prior Size | “Expected n” | “Likely Range” |
|---|---|---|---|
| 0.5 | -- | 100 | 84, 117 |
| 0.5 | 50 | 100 | 77, 135 |
| 0.5 | 10 | 100 | 66, 188 |

Based on 10,000 simulations

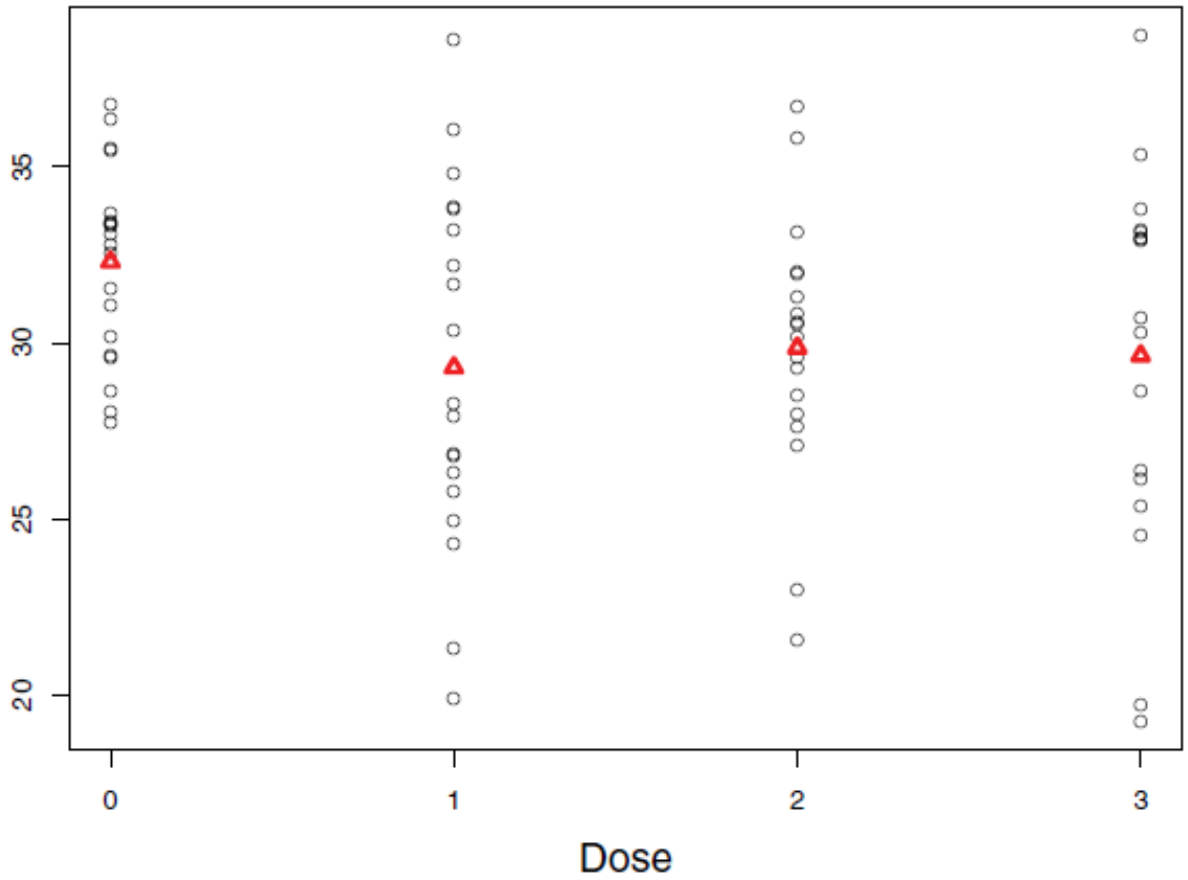
# Effect on Recruitment Time

- Add effect of uncertainty in recruitment rate on likely range of recruitment times

| variance | α | β | Prevalence | Prior Size | "expected time" | "Likely Range" |
|---|---|---|---|---|---|---|
| -- | -- | -- | 0.5 | -- | 10 | 7.8, 12.4 |
| 1 | 100 | 10 | 0.5 | -- | 10.1 | 7.2, 13.0 |
| 1 | 100 | 10 | 0.5 | 50 | 10.1 | 7.0, 14.6 |
| 1 | 100 | 10 | 0.5 | 10 | 10.1 | 6.1, 19.5 |
| 2 | 50 | 5 | 0.5 | -- | 10.2 | 6.8, 13.6 |
| 2 | 50 | 5 | 0.5 | 50 | 10.1 | 6.8, 15.2 |
| 2 | 50 | 5 | 0.5 | 10 | 10.1 | 6.0, 19.6 |
| 5 | 20 | 2 | 0.5 | -- | 10.5 | 5.7, 15.3 |
| 5 | 20 | 2 | 0.5 | 50 | 10.2 | 6.3, 16.9 |
| 5 | 20 | 2 | 0.5 | 10 | 10.4 | 5.7, 22.0 |

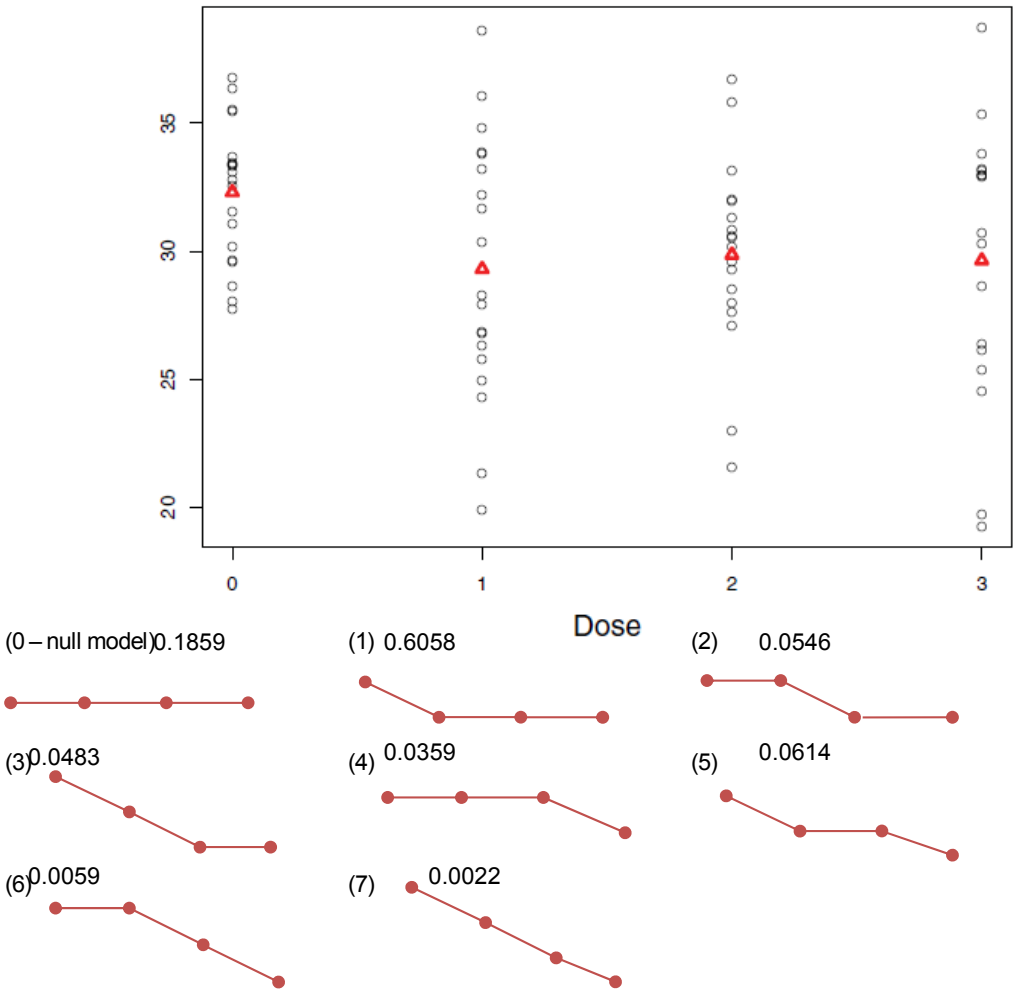
Based on 10,000 simulations

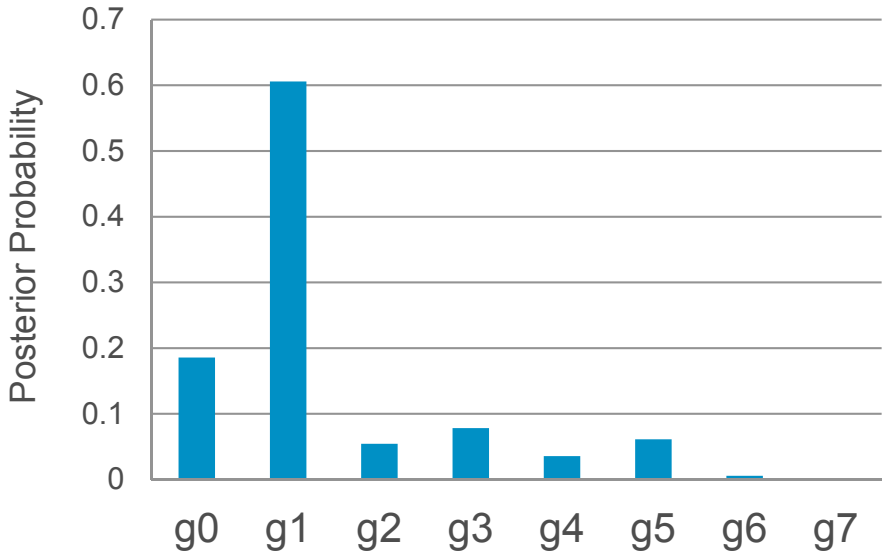
# Early Biomarkers
## Example




Otava M., Shkedy Z., Lin D., Göhlmann H.W.H., Bijnens L., Talloen W., Kasim A. (2014). Dose–Response Modeling Under Simple Order Restrictions Using Bayesian Variable Selection Methods. *Statistics in Biopharmaceutical Research*, 6:3, 252-262.
Otava M. (2014). Bayesian variable selection in dose-response relationship concept. International Biometric Conference, Florence.
Otava M. (2013). Bayesian Variable Selection Method for Modeling Dose-Response Microarray Data Under Simple Order Restrictions. Bayes2013, Rotterdam.

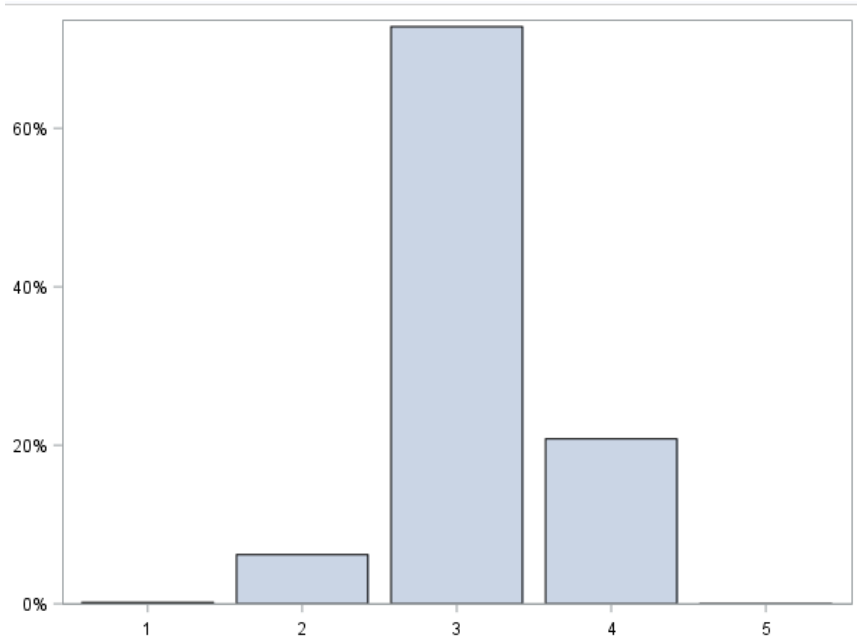
# Early Biomarkers
## Example - Results



# Biomarkers
## Bayesian Cutpoints

- Often, very limited data to determine cutpoint(s)
- Increasingly necessary to determine target population
  - So: early in the process
- Use of Bayesian techniques
  - Use of preclinical data / literature data?
  - Investigation in IDEAS



# Summary and Discussion

- Increased used of advanced statistical methods in research and in early clinical development
  - Increasing use of Bayesian methodology
  - Facilitate translation of knowledge (?)
  - High need to invest in “loop back”
- High level of interaction needed
  (specification of questions, determination of priors, … )
- Highly interdisciplinary
  - “mathematical functions”
  - Clinical and preclinical functions

**The business of the statistician is to catalyze the scientific learning process.**

- George Box

Science For A Better Life

# Thank you!