Tutorial 3
SUPPORTING THE PATHWAY TO TRIALS FOR RARE DISEASES: CLINICAL TRIAL DESIGN AND OTHER CONSIDERATIONS

Part 1
Clinical trial design for small sample population group trials

Ralf-Dieter Hilgers, on behalf of the IDeAl Project Team,
IDeAl coordinator, Department of Medical Statistics, RWTH Aachen

General Comments

Outline

In this lecture, we will:

- Discuss the pathway from the statistical perspective
- Describe some problems of the pathway
- Describe some points of improvement for the pathway with special interest in small population groups (SPG)

Overview: Orphan Drug Use in Rare Disease

- Transfer from animals to human
- Randomization
- Outcomes – scores
  - Long term survival
  - Identify biomarker
- Design
- Building computational models of a disease and simulate trial designs
- Extrapolation
- Mirror clinical pathways

Animal Experiments

OPEN ACCESS Freely available online

PLoS MEDICINE

Research in Translation

Can Animal Models of Disease Reliably Inform Human Studies?

H. Bart van der Worp[1]*, David W. Howells[2], Emily S. Sena[2,3], Michelle J. Porritt[2], Sarah Rewell[2], Victoria O'Collins[2], Malcolm R. Macleod[3]

1 Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, Utrecht, The Netherlands, **2** National Stroke Research Institute & University of Melbourne Department of Medicine, Austin Health, Melbourne, Australia, **3** Department of Clinical Neurosciences, University of Edinburgh, Edinburgh, United Kingdom

Methodological Flaws

- Reduced Internal Validity
  - Randomization
  - Blinding
  - Sample Size Calculation
  - Control Physiological Variables / disease model
  - Statistical Methods
- Reduced External Validity
  - Critical disparities between animal model and clinical trials
  - Publication Bias

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button[1,2], John P. A. Ioannidis[3], Claire Mokrysz[1], Brian A. Nosek[4], Jonathan Flint[5], Emma S. J. Robinson[6] and Marcus R. Munafò[1]

- Low probability of finding true effects
- Low positive predictive value when an effect is claimed
- Exaggerated estimate of the magnitude of the effect when a true effect is discovered
  - Winner's curse
  - Proteus phenomenon

Nature Reviews Neuroscience | AOP, published online 10 April 2013; doi:10.1038/nrn3475

What is Underpowered

- Significance level 5%, Power 80%
- $P_{standard}$ = 30%, P_{new} = 60 %, **δ=30%**
  Necessary sample size 49 /group
- Now only 20 / group could be enrolled
- $P_{standard}$ = 30%, P_{new} = 60 %, **δ=30%**

the power is 34%

[Calculation by: nQuery 7.0 PTT1-1]

General Comments on Clinical Trials

Recommendation

Three basic requirements for any clinical trial:

1. Trial should examine an important research question
2. Trial should use a rigorous methodology that can answer the question of interest
   - often problematic in small trials
3. Trial must be based on ethical considerations and assure that risks to subjects are minimized

Basic Concepts of Clinical Trials

- Experiments
  - Statement of problem
  - Objective of the study
  - Choice of response variable
  - Selection of factors to be varied
  - Choice of levels of these factors
- Design
  - Sample Size
  - Method of randomization
  - Mathematical Model
  - Hypothesis
  - Blind Procedure

Randomization

Randomization

… allocation of study units to receive one of the study treatments by randomization ("tossing a coin")

… tends …

- to produce study groups comparable with respect to known risk factors,
- *may help to remove bias*
- *guarantees that statistical tests will have valid significance levels*

What is Bias – Only Ivory Tower

Worse prognosis $P_{s,W} = 10\%$, $P_{n,W} = 30\%$

Better prognosis $P_{s,B} = 40\%$, $P_{n,B} = 60\%$

"mean" patients: $P_s = 25\%$, $P_n = 45\ \%$

[n=98/group]

Selection:

- **$P_{s,B}$ vs $P_{n,W}$: 40% vs 30% [Power 25%]**
- 80 % of "better" would be allocated to S
  80 % of "worse" would be allocated to N
  ➔ 17% under S vs 18% under N
  ➔ Power 2% in 98/group trial

[Calculation by: nQuery 7.0 PTT1-1]

Randomization

- Bias may cause serious problems
  - Patients with better prognosis could be preferred by predictions
  - Prolonged enrolment could cause time trend bias
- Balancing “cofactors” is necessary to prove treatment effect
- Some randomization procedures (allocation schemes) bear the risk of more bias

Randomization – Selection Bias

Worse prognosis $P_{s,W} = 10\%$, $P_{n,W} = 30\%$

Better prognosis $P_{s,B} = 40\%$, $P_{n,B} = 60\%$

δ=30%

- 12 Patients:
- SSSSSS NNNNNN
  - .25+.40+.40+.40+.40+.40=**.375**
  - .60+.60+.60+.60+.60+.60=**.60**

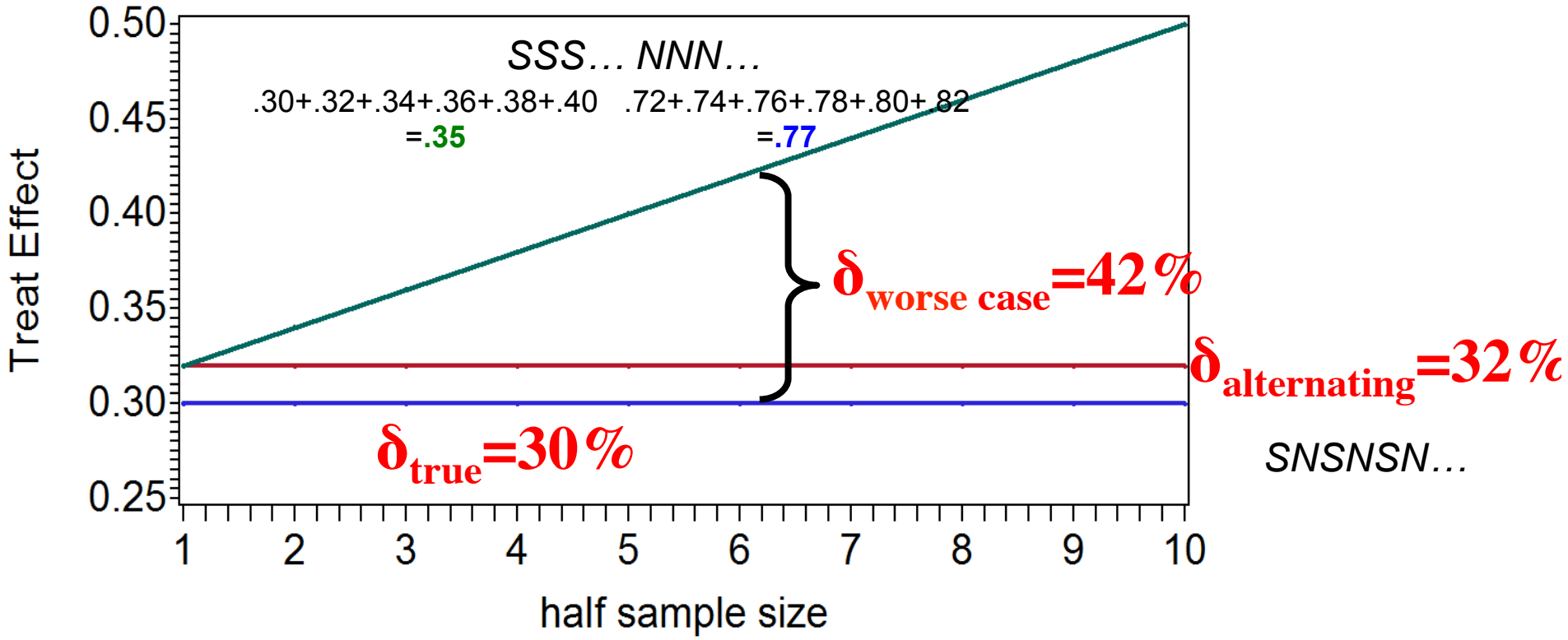
δ=22.5%

- SNSNSNSNSNSN
  - .25+.25+.25+.25+.25+.25=**.25**
  - .60+.60+.60+.60+.60+.60=**.60**

δ=35%

Randomsierung – Time Trend Bias

- 2% learning effekt; $P_{standard}$ = .30, P_{new} = .60



Randomization

Aim: Recommendation minimize risk of bias

- Fixed allocation randomization
  - Simple randomization
  - Blocked randomization
  - "Urn" randomization
- Adaptive randomization
  - Baseline adaptive randomization
  - Response adaptive randomization

Combine with stratification

Kennes, Cramer, Hilgers, Heussen (2011)
Tamm, Cramer, Kennes, Heussen (2011)
Kennes, Cramer, Hilgers, Heussen (submitted)
Tamm, Hilgers, Rosenberger (submitted)

Endpoints vs Surrogate Endpoints

Endpoints in SPG

- Conventional clinical trials often require hundreds of patients and take years to complete
- This is particularly problematic in small population diseases the available number of potential study participants is small Long trials should be avoided to reduce the probability of study drop-out (which further depletes the data)
- The required time and sample size of a clinical trial is strongly affected by the endpoint that is used.

Surrogate endpoints in SPG

- **True endpoint (T):** the best indicator of the therapeutic response
- **Surrogate endpoint (S):** could be used instead of T. S is 'easier' to measure and allows for an accurate prediction of T and of the treatment effect on T
- A 'good' S has two main properties (Buyse 2000)
  - **individual-level surrogacy:** *S should allow for a good prediction of T*
  - **trial-level surrogacy:** *The treatment effect on S should allow for a good prediction of the treatment effect on T*

Surrogate endpoints evaluation: example

- Setting:

  A clinical trial in age-related macular degeneration (ARMD)

  N=181 patients from 36 centers participated

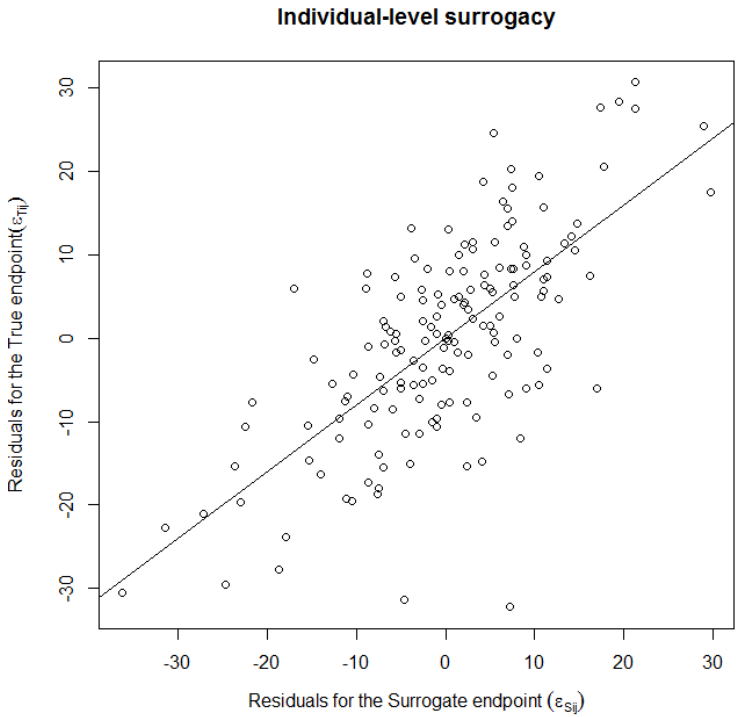
  Endpoints: change in visual acuity measured after

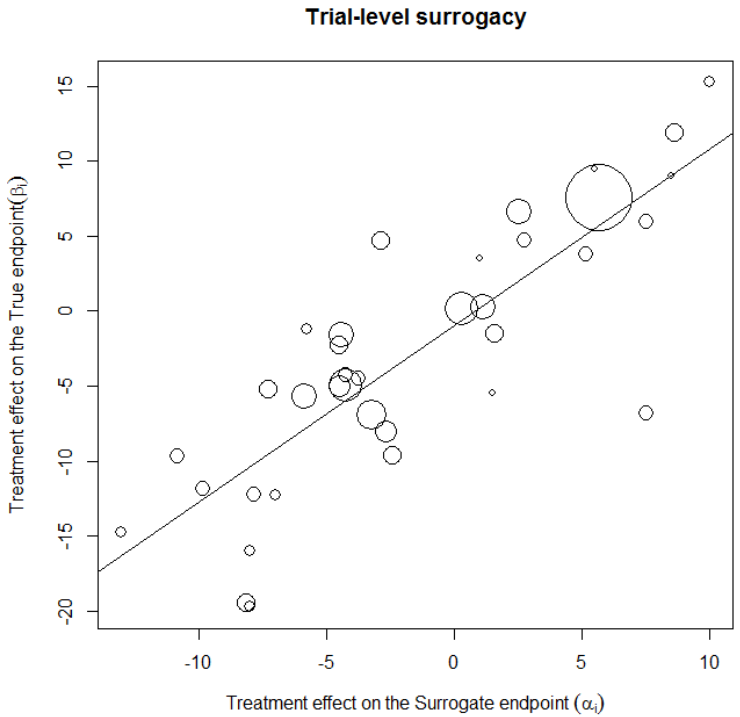
  - 6 months (the candidate surrogate)
  - 12 months (the true endpoint)

  Treatment: Inferon-α and placebo

- Question: is change in visual acuity after 6 months a good surrogate (S) for change in visual acuity after 12 months (T)
- Center is used as the clustering level
- The freely available R package Surrogate is used to analyze the data

Surrogate endpoints evaluation: example





$$R^2_{indiv} = 0.487$$

$$95\%\ CI : [0.381, 0.592]$$

$$R^2_{trial} = 0.703$$

$$95\%\ CI : [0.533, 0.872]$$

Surrogate endpoints in SPG

- The formal evaluation of a surrogate endpoint (i.e., the quantification of individual- and trial-level surrogacy) is not a trivial endeavor
- Especially when data are sparse, problems may occur (e.g., the statistical models needed to estimate individual- and trial-level surrogacy may not converge)

Pharmacogentics
Repeated / Crossover

Statistical Statements to Individual Response

- Personalized medicine is related to individual response to treatment
- Individual Response is usually not identifiable in clinical trials
- Individual response to treatment can be estimated by subject by treatment interaction
- Different Designs are necessary

Crossover / Repeated measures design

A cross-over trial is one in which subjects are given sequences of treatments with the objective of studying differences between individual treatments (or sub-sequences of treatments).

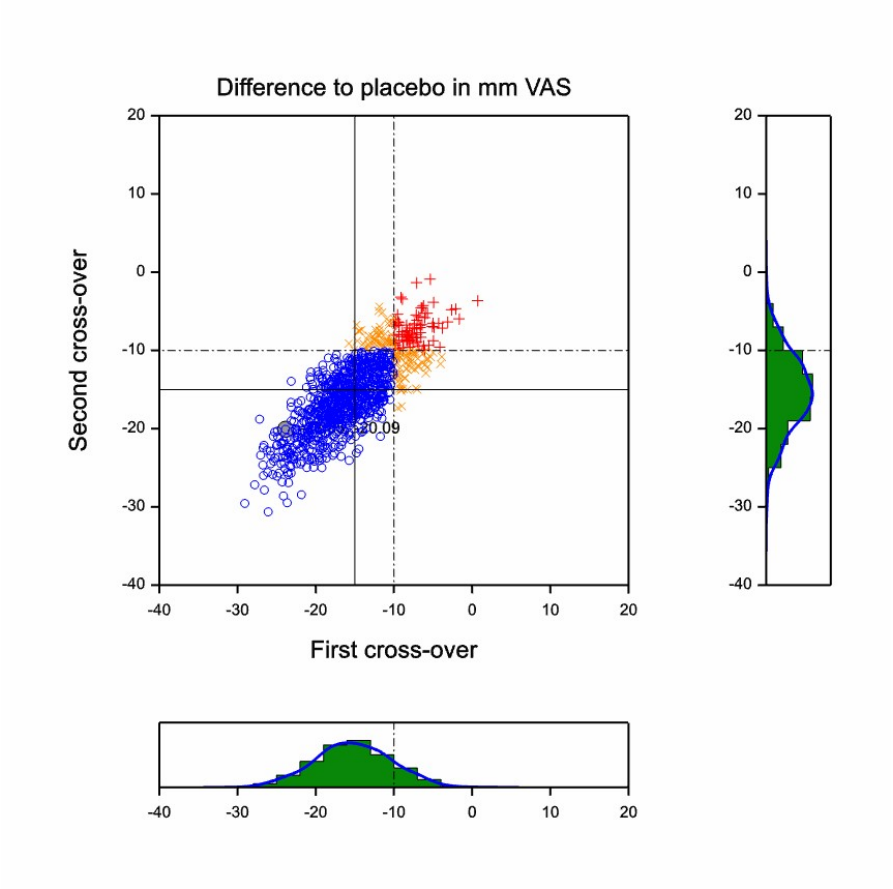
Senn, 1993, Cross-over Trials in Clinical Research

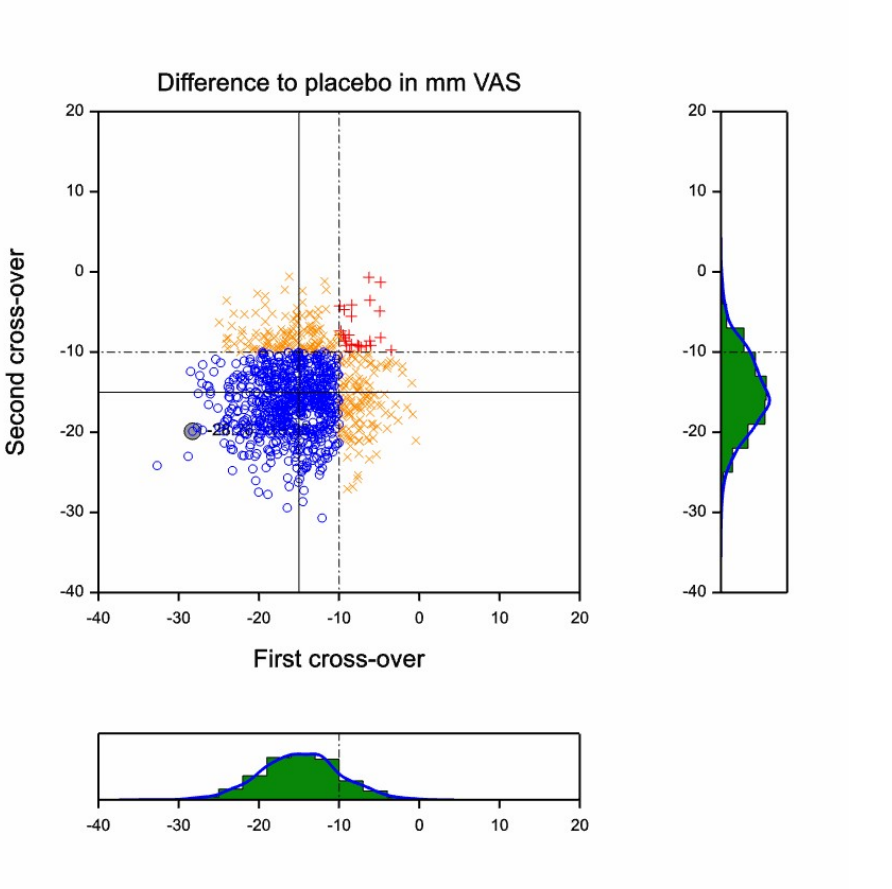
A Repeated measures design involves multiple observations or response variables for each subject.

- Repeated measurements over time (longitudinal)
- Multiple measurements on same subject

A Thought Experiments - Design

	First Cross-over		Second Cross-over	
	Period			
Sequence	1	2	3	4
I	A	P	A	P
II	P	A	P	A
III	A	P	P	A
IV	P	A	A	P


Difference to placebo in mm VAS
Second cross-over
First cross-over
-20.09


Difference to placebo in mm VAS
Second cross-over
First cross-over

Statistical Statements to Individual Response

- The (marginal) distribution does not inform about the prediction of responders
- Responding to a treatment depends on the correlation of repeated treatment applications
- Conventional clinical trials provide averaged treatment effects – not individual responding
- Repeated Crossover trial enables estimation of patient by treatment interaction
- Other designs are necessary

Adaptive Design

Adaptive (flexible) vs Fixed Designs

Classical frequentist trials

- details of design and analysis must be prefixed in advance (population, treatments, doses, main and secondary outcome variable(s), analysis strategy, sample sizes,...)
- lack of flexibility to react to information from inside or outside the trial

Flexible (adaptive) design

- allow for mid-trial design modifications based on all internal and external information gathered at interim analyses without compromising the type I error rate
- To control the type I error rate, the design modifications need **not** be specified in advance.

Popular types of adaptations

- Sample size reassessment
  - Based on nuisance parameter estimates
  - Based on effect estimates
- Selection of treatments/doses
  - Adaptive Seamless Designs
- Selection of subgroups
  - Adaptive Enrichment Designs
- Adding or dropping of interim analyses
- Modification of endpoints

. . .

Writing amendments for "online" design modifications will not be the general solution!

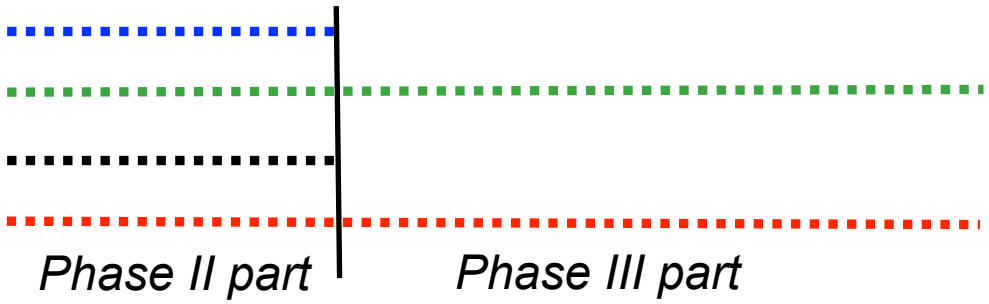
ADAPTIVE DESIGNS (IF CAREFULLY PLANNED AND CONDUCTED) GUARANTEE A STRICT TYPE I ERROR CONTROL IN CASE OF DESIGN MODIFICATIONS IN ON-GOING CLINICAL TRIALS

Separate Phase II and III Trials

- Conduct phase II trial
- Plan phase III trial based on the information from phase II trial (which dose, which subgroup, which number of patients etc …
- Conduct confirmatory phase III trial.
  Demonstrate efficacy using only phase III trial data.
  INEFFICIENT IN SMALL POPULATIONS!

Adaptive (seamless) phase II +III designs

Learning (Phase II)



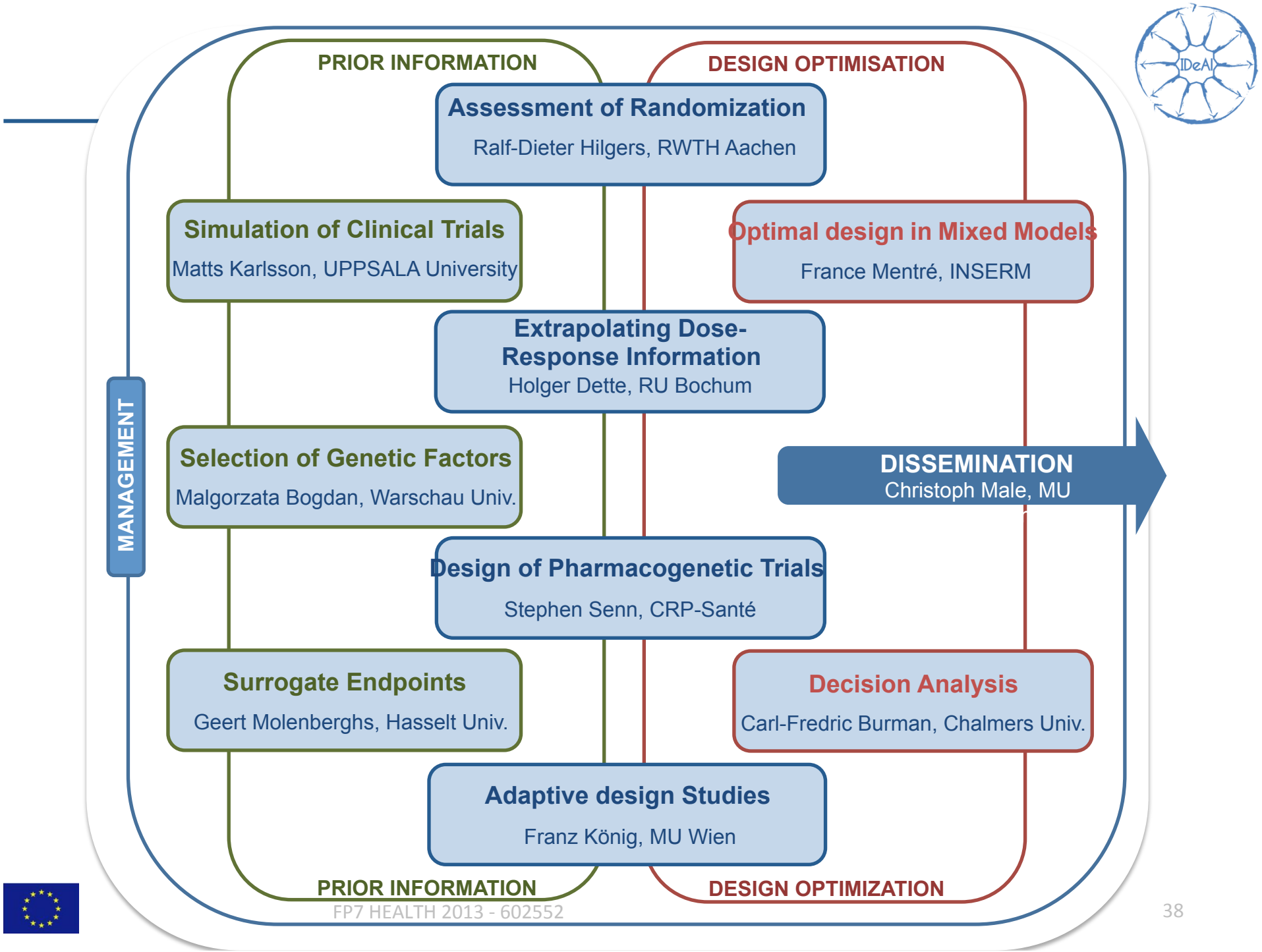
- Conduct phase II as internal part of a combined trial
- Plan phase III part based on data from phase II part
- Conduct phase III trial as internal part of the same trial
- Demonstrate efficacy using ALL data from phase II+III.

Potential benefits of Confirmatory Adaptive Designs

- formally integrate learning & confirming aspects in a trial
- can use all accumulated data to perform design adaptations (as sample size reassessment, selecting of groups, …)
- smaller time lag between phase II and phase III. Speeds up the drug development process.
- allows **USE OF ALL AVAILABLE DATA** for decision making (efficacy hypothesis testing). This saves resources (patients), costs and time. VERY IMPORTANT IN SMALL POPULATIONS!
- may spare the preparation time for a second trial (only one protocol needed, one approval by ethics committee,..
- allow to react flexibly to unexpected events

"Such a design has the potential to speed up the process of drug development or can be used to allocate resources more efficiently without lowering scientific and regulatory standards."
EMA REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN (2007).

Conclusion - Remarks


PRIOR INFORMATION
DESIGN OPTIMISATION
Assessment of Randomization
Ralf-Dieter Hilgers, RWTH Aachen
Simulation of Clinical Trials
Matts Karlsson, UPPSALA University
Optimal design in Mixed Models
France Mentré, INSERM
Extrapolating Dose-Response Information
Holger Dette, RU Bochum
MANAGEMENT
Selection of Genetic Factors
Malgorzata Bogdan, Warschau Univ.
DISSEMINATION
Christoph Male, MU
Design of Pharmacogenetic Trials
Stephen Senn, CRP-Santé
Surrogate Endpoints
Geert Molenberghs, Hasselt Univ.
Decision Analysis
Carl-Fredric Burman, Chalmers Univ.
Adaptive design Studies
Franz König, MU Wien
PRIOR INFORMATION
DESIGN OPTIMIZATION

HOW TO STAY IN CONTACT WITH IDEAL

- IDEAL: Intgrated Design and Analysis of Small Population Group Trials

- VISIT THE **IDEAL WEBPAGE**
  - http://www.ideal.rwth-aachen.de
- Get **LinkedIn** IDEAL – FP7 Project
  - http://www.linkedin.com/groups/IDEAL-FP7-Project-6556030
- **Twitter** @ideal_fp7
  - https://twitter.com/ideal_fp7

Summary

- Animal experiments lack of external and internal validity, which hampers translational medicine
- Small trials need special methodologies
- Transfer to personalised medicine is not straightforward
- IDeAl will develop new statistical methodologies in form of Integrated Designs and Analysis for small population group trials