Interim Analysis, Adaptations and Master Protocols

2024-04-23 EJP RD_WP20 Intermediate Course

Franz König und Martin Posch

Center for Medical Data Science, franz.koenig@meduniwien.ac.at

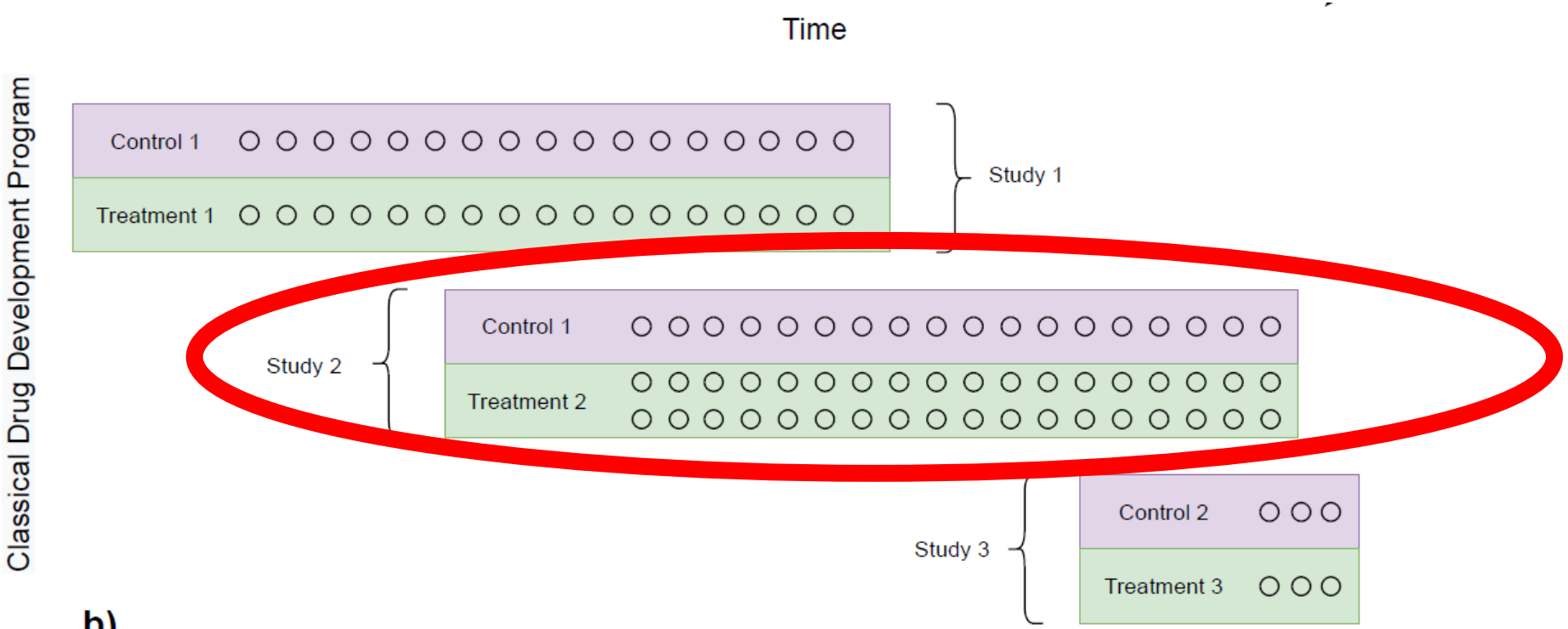
https://data-science.meduniwien.ac.at/

Content

- Introduction – standard development programs and clinical trials
- Interim Analysis and Group Sequential Designs
- Adaptive designs
- Adaptive Seamless Trials
- Master Protocols
- AOB

Classical Drug Development Programs

Meyer et al. (2020b)



Traditionally: type I error control at study level, regardless which other studies are performed

Type I error:
Error of rejecting a null hypothesis when it is acutally true

Want to learn from any new data when they become available – need for interim analysis!

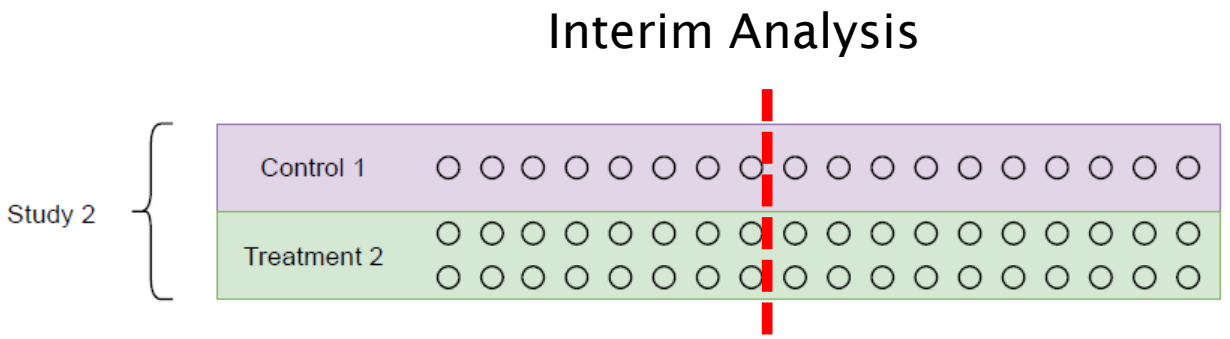
Traditionally: Frequentist hypotheses tests for decision making in confirmatory clinical trial

- Traditional decision making in confirmatory clinical trials is based on hypothesis testing
- The null hypothesis **"The experimental treatment is not superior to control"** is tested with a statistical test
- Based on the clinical trial data a **p-value** is calculated
- If **$p<0.05$** the null hypothesis is rejected and the drug is declared efficacious
- This guarantees that the probability **of a false positive result** (given the treatment does not work) **is lower than 5%**
- However, if **multiple tests** are performed with the same threshold of 0.05, the risk of at **least one false positive conclusion increases**

> "Control of the study-wise rate of false positive conclusions at an acceptable level α is an important principle and is often of great value in the assessment of the results of confirmatory clinical trials."

Points to consider on multiplicity issues in clinical trials, EMA (2002)

Interim Analysis



- Definition:

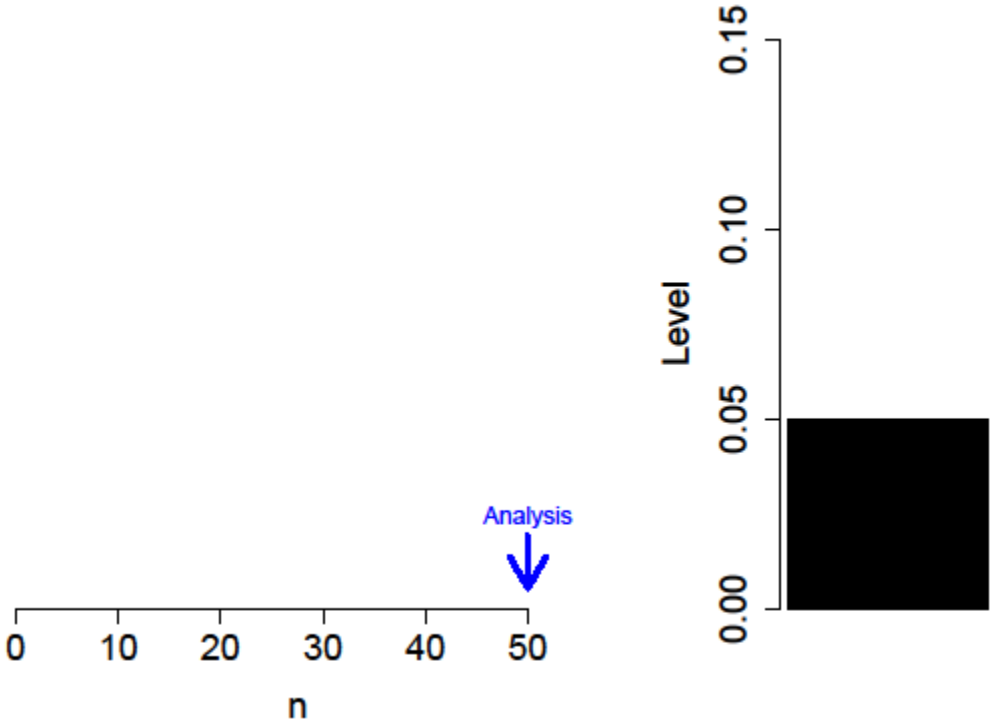
Any analysis, summary, or inspection of unblinded trial data during an ongoing trial

- Motivation:
  - Ethical
  - Economical
  - Possibility to stop the trial early for
    - safety concerns,
    - overwhelming effect, or
    - futility (= lack of efficacy)
  - Possibility to modify design aspects of the ongoing trial
  - Use observed information to plan future studies

What will happen if we ignore the pre-planned fixed sample design and use the convential level alpha test to decide whether to stop or continue a trial?

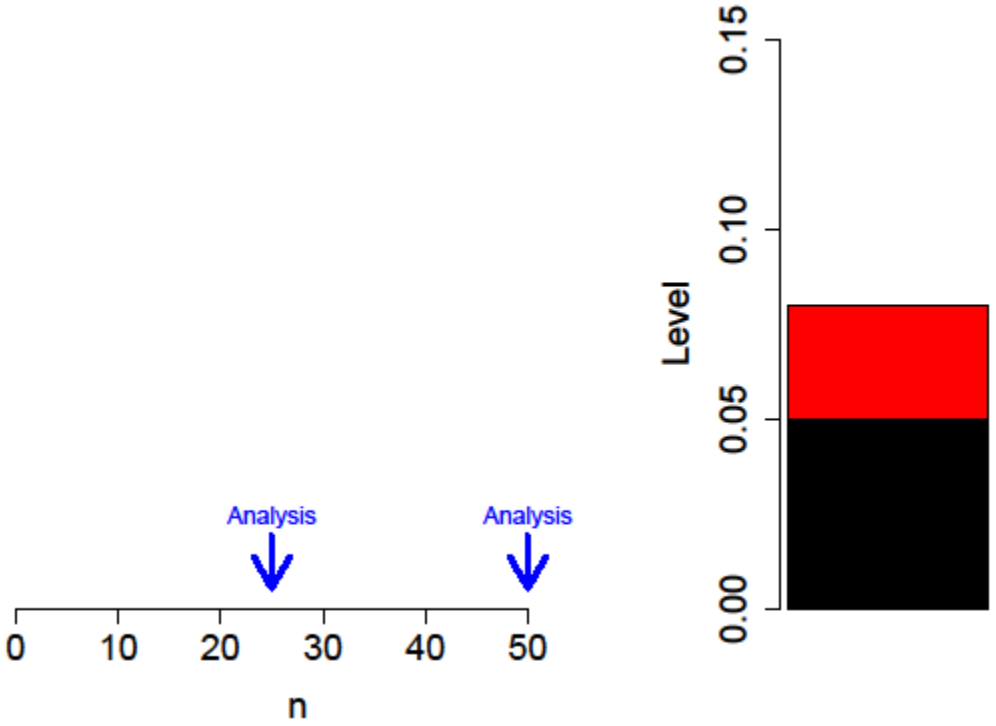
Hunting for Significance Tests for a Single Hypothesis

1 Analysis at the level α=0.05



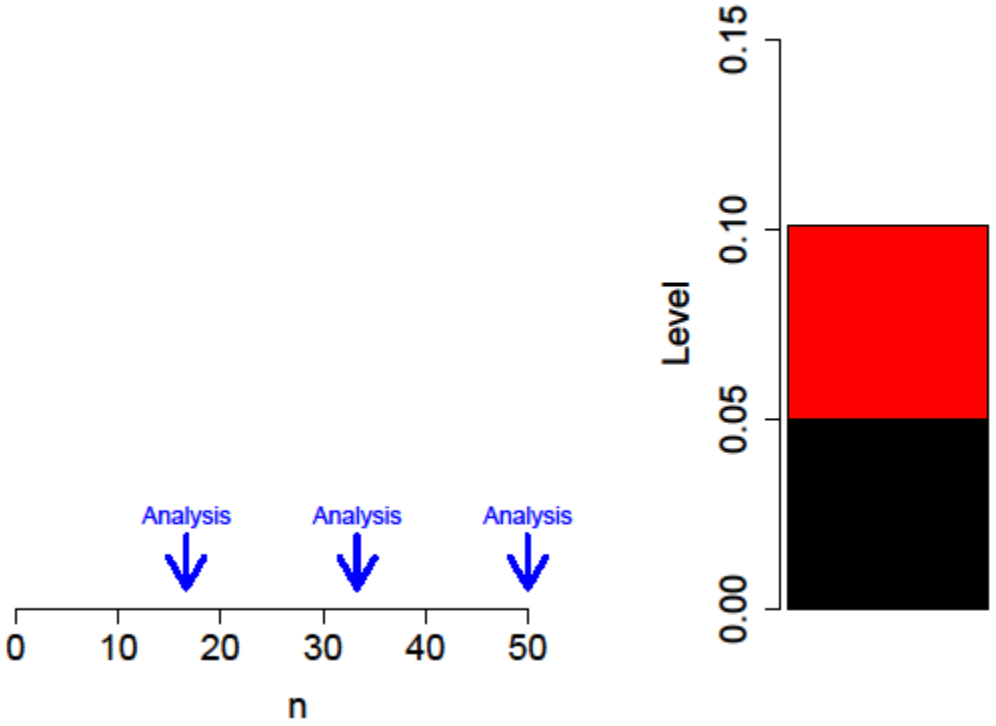
Hunting for Significance Tests for a Single Hypothesis

2 Analyses at the level $\alpha=0.05$



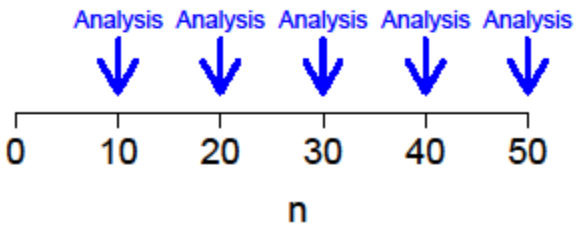
Hunting for Significance Tests for a Single Hypothesis

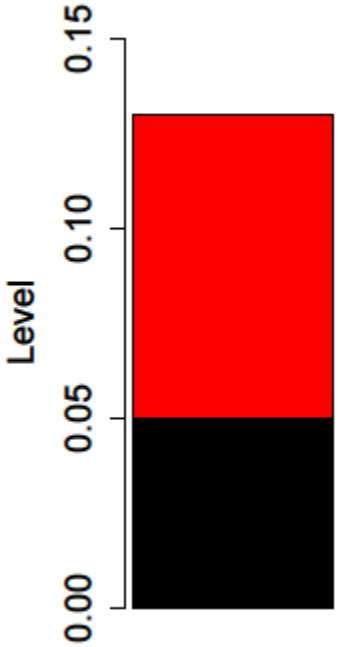
3 Analyses at the level $\alpha=0.05$



Hunting for Significance Tests for a Single Hypothesis

5 analysis at the level α=0.05





(see Armitage 1969)

Key Issue

Type I Error Rate Control

- Group Sequential Designs
  - Design the trial with multiple analyses over the course of the trial. Predefine a set of critical values (boundaries) to account for the multiplicity
  - The overall boundary crossing probability (or the Type I error rate) is maintained at a pre-specified significance level α
  - Typical group sequential design include Pocock (1977), O´Brien and Fleming (OBF, 1979) or designs with **alpha spending function Lan & Demets** (1983)
- Adaptive Designs
  - Designs allow design modifications during the trial (e.g., randomization fraction, sample size, treatment selection, etc.) based on IA results of current trial or from external resources
  - Describe decision rules/criteria in protocol, i.e., changes are made "by design" and not "ad hoc"
  - Special analysis methods needed, e.g., based on adaptive combination tests or conditional error function

E.g. when planning for 4 interim analyses (IA) (two-sided tests)

instead of

$p<0.05$

one has to use

Pocock

$p<0.0158$

O´Brien & Fleming

$p<0.0001$ (1st IA)
$p<0.0013$ (2nd IA)
$p<0.0084$ (3rd IA)
$p<0.0226$ (4th IA)
$p<0.0413$ (final analysis)

shiny.rpact.com

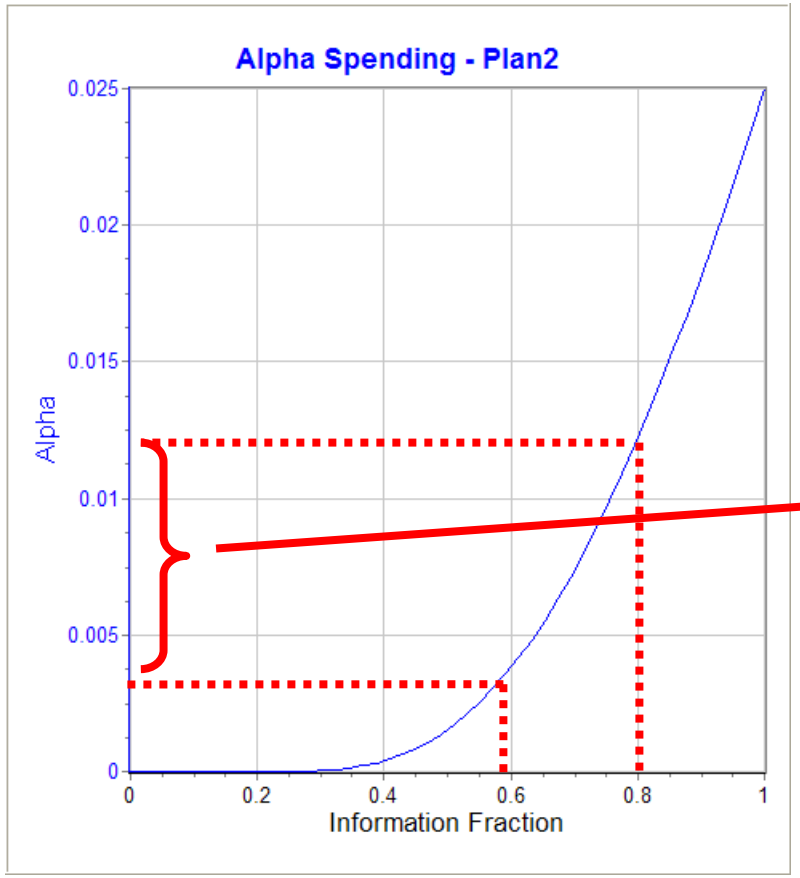
Some remarks for group seqeuntial designs

- Stopping boundaries, number and sample sizes of interim looks have to be pre-specified in advance.
- The first group sequential designs were originally designed for equally spaced interim looks only
- In clinical trials one might not exactly achieve the pre-planned stagewise sample size
  - e.g., interim looks might be determined by calendar time.
  - What is the impact on the type I error in case of (small) deviations?
- Mainly interested in demonstrating superiority of the new treatment against control (and not vice-versa).
  - ➔ one-sided in tests, e.g., (assuming larger is better)

$$H_0: \mu_T \leq \mu_C \text{ vs } H_1: \mu_T > \mu_C$$

  - Different stopping boundaries for superiority and inferiority (futility stop)

Alpha Spending Function Lan&DeMets 1983



- Pre-specify alpha spending function and maximum sample size (=information fraction=1)
- Stopping boundaries as continuous function of accrual
- Timing and number of interim analysis are flexible. This allows unequally spaced interim looks (unequal group sample sizes).
- The increment in the alpha spending function between two interim analysis is used to determine the critical boundaries.
- However, the adaptations (number and timing of interim analyses) may not depend on the interim data. (One could use some odd alpha spending functions to inflate the actual type I error rate.)

$H_0: \mu_T \leq \mu_C$ vs $H_1: \mu_T > \mu_C$

How to plan and compare group sequential designs?

- Power and ASN under H_0 and different alternatives
- Stopping probabilities (efficacy/futility) per stage
- Number and timing of interim analysis

- Minimum sample size
  - The decision to stop (efficacy/superiority or futility) will be based on the primary endpoint only.
  - Is the first stage sample size large enough to check consistency in important other variables and/or subgroups?

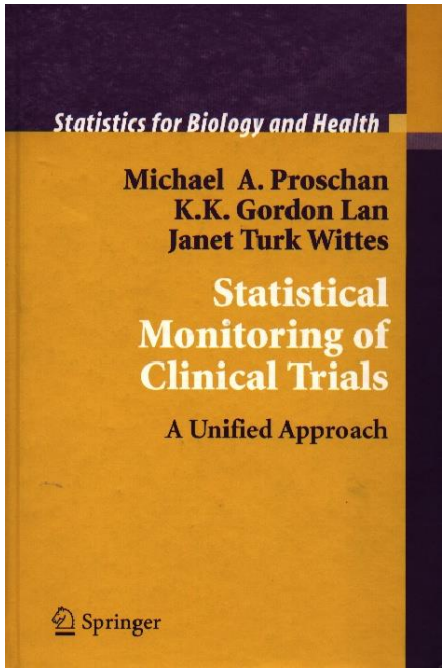
- Check which effect would be required for an early rejection.
  - Can such a large effect be expected?

- Maximum total sample size

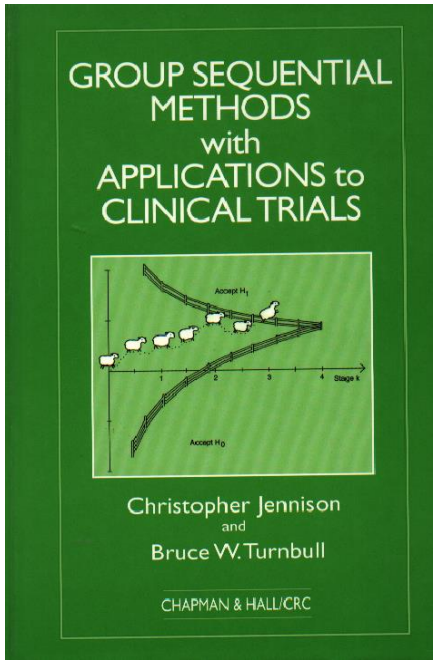
- Binding/non-binding stopping for futility boundaries

- How shall other (key) secondary endpoints be tested?

Some references

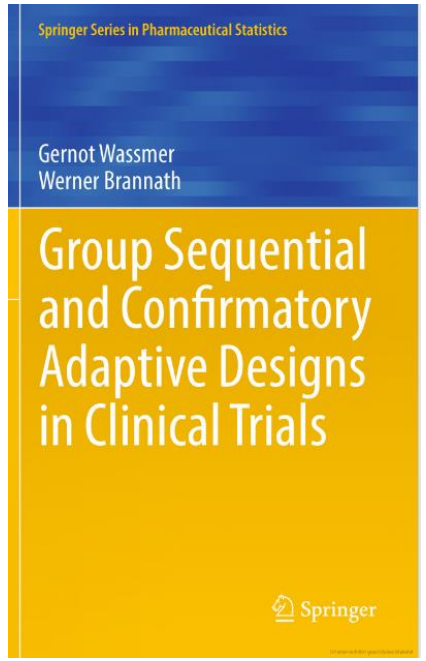


Springer, 2006



Chapman & Hall, 2000

Cologne, 1999



Springer, 2016

If more flexibility is needed …
… then use adaptive designs

Classical vs Flexible (Adaptive) Trials

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.

Some History of Adaptive Designs

1989	Bauer: “Multistage Testing with Adaptive Designs”
1995	Proschan & Hunsberger: “Designed Extension of Studies Based on Conditional Power”
2007	EMA Reflection Paper
2010	FDA Draft Guidance (Drugs and Biologics)
2015	FDA Draft Guidance (Devices, CDRH, CBER) - finalized 2017
2018	New FDA Draft Guidance (Drugs and Biologics) – finalized 2019
2019	Concept paper ICH E20 Adaptive Clinical Trials

Featured Article

Statistics in Medicine

Received 2 September 2014, Accepted 19 February 2015 Published online 16 March 2015 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6472

Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls

Peter Bauer,[a] Frank Bretz,[b,c] Vladimir Dragalin,[d] Franz König[a] and Gernot Wassmer[e,f*†]

http://dx.doi.org/10.1002/sim.6472 (Open Access)
With invited discussion by Hung, Wang and Lawrence; Mehta and Liu; Vollmar; Maurer

Some History of A

1989	Bauer: "M
1995	Proschan Condition
2007	EMA Refle
2010	FDA Draft
2015	FDA Draft
2018	New FDA
2019	Concept p

Elsäßer et al. Trials 2014, 15:383
http://www.trialsjournal.com/content/15/1/383

TRIALS

RESEARCH Open Access

Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency

Amelie Elsäßer[1], Jan Regnstrom[2], Thorsten Vetter[2], Franz Koenig[3], Robert James Hemmings[4], Martina Greco[2], Marisa Papaluca-Amati[2] and Martin Posch[3*]

Abstract

Background: Since the first methodological publications on adaptive study design approaches in the 1990s, the application of these approaches in drug development has raised increasing interest among academia, industry and regulators. The European Medicines Agency (EMA) as well as the Food and Drug Administration (FDA) have published guidance documents addressing the potentials and limitations of adaptive designs in the regulatory context. Since there is limited experience in the implementation and interpretation of adaptive clinical trials, early interaction with regulators is recommended. The EMA offers such interactions through scientific advice and protocol assistance procedures.

Methods: We performed a text search of scientific advice letters issued between 1 January 2007 and 8 May 2012 that contained relevant key terms. Letters containing questions related to adaptive clinical trials in phases II or III were selected for further analysis. From the selected letters, important characteristics of the proposed design and it context in the drug development program, as well as the responses of the Committee for Human Medicinal Products (CHMP)/Scientific Advice Working Party (SAWP), were extracted and categorized. For 41 more recent procedures (1 January 2009 to 8 May 2012), additional details of the trial design and the CHMP/SAWP responses were assessed. In addition, case studies are presented as examples.

Results: Over a range of 5½ years, 59 scientific advices were identified that address adaptive study designs in phas II and phase III clinical trials. Almost all were proposed as confirmatory phase III or phase II/III studies. The most frequently proposed adaptation was sample size reassessment, followed by dropping of treatment arms and population enrichment. While 12 (20%) of the 59 proposals for an adaptive clinical trial were not accepted, the great majority of proposals were accepted (15, 25%) or conditionally accepted (32, 54%). In the more recent 41 procedures, the most frequent concerns raised by CHMP/SAWP were insufficient justifications of the adaptation strategy, type I error rate control and bias.

Conclusions: For the majority of proposed adaptive clinical trials, an overall positive opinion was given albeit with critical comments. Type I error rate control, bias and the justification of the design are common issues raised by th CHMP/SAWP.

Keywords: adaptive design, EMA, FDA, seamless designs, scientific advice, clinical trials, phase II-III, pivotal trial, orphan drugs

* Correspondence: Martin.Posch@meduniwien.ac.at
[3]Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria
Full list of author information is available at the end of the article

Collignon et al. Trials (2018) 19:642
https://doi.org/10.1186/s13063-018-3012-x

Trials

RESEARCH Open Access

Adaptive designs in clinical trials: from scientific advice to marketing authorisation to the European Medicine Agency

Olivier Collignon[1,2*], Franz Koenig[3], Armin Koch[4], Robert James Hemmings[5], Frank Pétavy[1], Agnès Saint-Raymond[1], Marisa Papaluca-Amati[1] and Martin Posch[3]

Abstract

Background: In recent years, experience on the application of adaptive designs in confirmatory clinical trials has accumulated. Although planning such trials comes at the cost of additional operational complexity, adaptive designs offer the benefit of flexibility to update trial design and objectives as data accrue. In 2007, the European Medicines Agency (EMA) provided guidance on confirmatory clinical trials with adaptive (or flexible) designs. In order to better understand how adaptive trials are implemented in practice and how they may impact medicine approval within the EMA centralised procedure, we followed on 59 medicines for which an adaptive clinical trial had been submitted to the EMA Scientific Advice (SA) and analysed previously in a dedicated EMA survey of scientific advice letters. We scrutinized in particular the submission of the corresponding medicines for a marketing authorisation application (MAA). We also discuss the current regulatory perspective as regards the implementation of adaptive designs in confirmatory clinical trials.

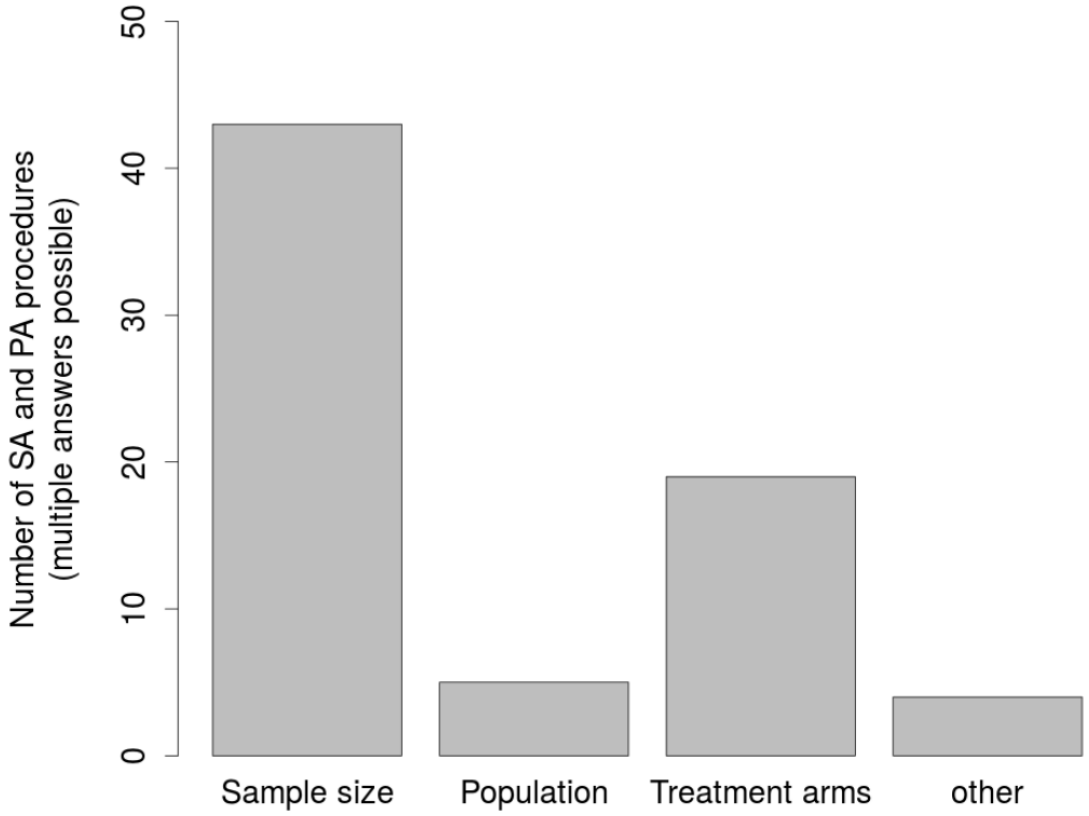
Methods: Using the internal EMA MAA database, the AdisInsight database and related trial registries, we analysed how many of these 59 trials actually started, the completion status, results, the time to trial start, the adaptive elements finally implemented after SA, their possible influence on the success of the trial and corresponding product approval.

Results: Overall 31 trials out of 59 (53%) were retrieved. Thirty of them (97%) have been started and 23 (74%) concluded. Nine of these trials (39% out of 23) demonstrated a significant treatment effect on their primary endpoint and 4 (17% out of 23) supported a marketing authorisation (MA). An additional two trials were stopped using pre-defined criteria for futility, efficiently identifying trials on which further resources should not be spent. Median time to trial start after SA letter was given by EMA was 5 months. In the investigated trial registries, at least 18 trial (58% of 31 retrieved trials) designs were implemented with adaptive elements, which were predominantly dose selection, sample size reassessment (SSR) and stopping for futility (SFF). Among the 11 completed trials including adaptive elements, 6 demonstrated a significant treatment effect on their primary endpoint (55%).

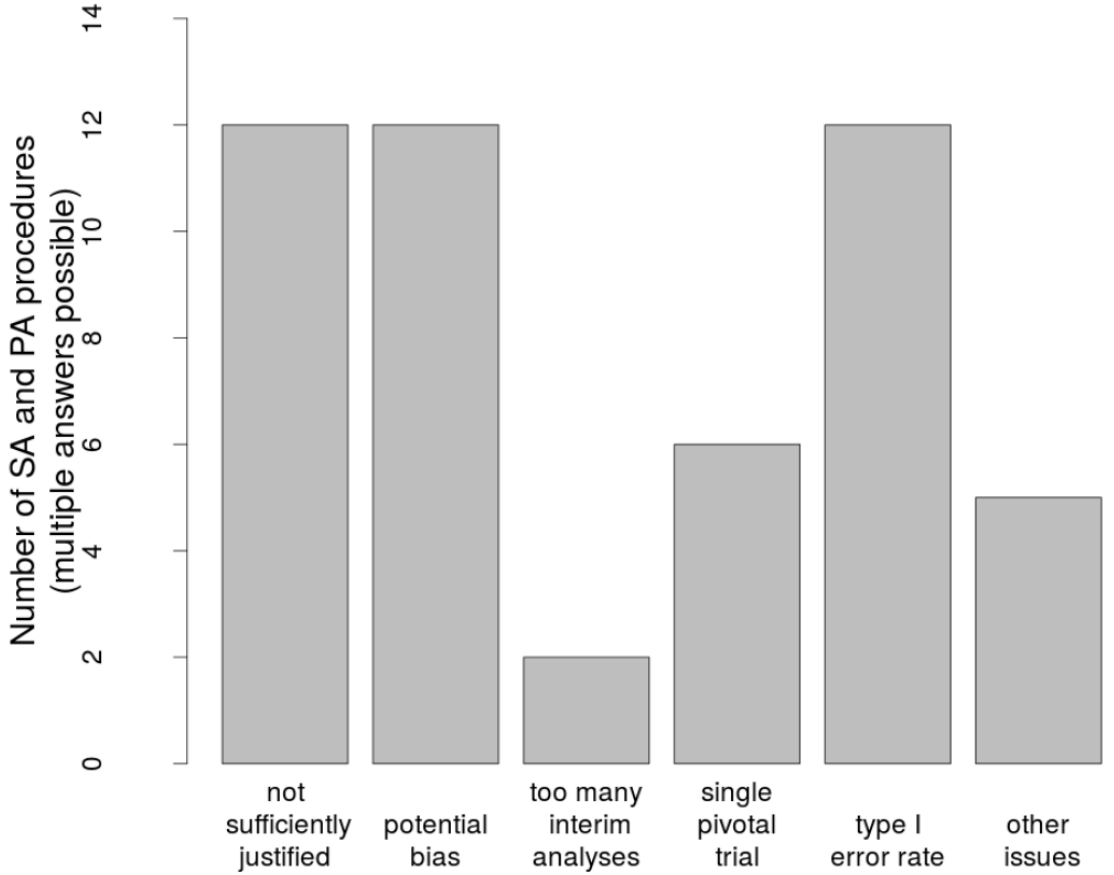
(Continued on next page)

* Correspondence: olivier.collignon@ema.europa.eu; olivier.collignon@lih.lu
[1]European Medicines Agency, 30 Churchill Place, London E14 5EU, UK
[2]Competence Center for Methodology and Statistics, Luxembourg Institute of Health, 1A-B, rue Thomas Edison, L-1445 Strassen, Luxembourg
Full list of author information is available at the end of the article

Which adaptions are of most interest in drug development?



What are the critical issues from a regulatory point of view?



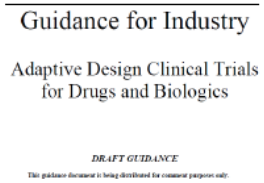
Elsässer, F. Koenig, ...M. Posch. Adaptive clinical trial designs for European marketing authorization: a survey of scientic advice letters from the European Medicines Agency. Trials 15, 383, (2014)
Collignon, O., Koenig, F., Koch... & Posch, M. (2018). Adaptive designs in clinical trials: from scientific advice to marketing authorisation to the European Medicine Agency. Trials, 19(1), 642.

Some regulatory definitions of adaptive designs


emea
European Medicines Agency
London, 18 October 2007
Doc. Ref. CHMP/EWP/2459/02
COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP)
REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN

A study design is called "adaptive" if statistical methodology allows the modification of a design element (e.g. sample-size, randomization ratio, number of treatment arms) at an interim analysis with full control of the type I error.

EMA 2007


Guidance for Industry

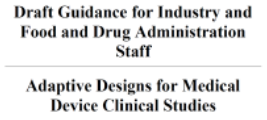
Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

A study that includes a prospectively planned opportunity for modification of one or more specified aspects of the study design and hypotheses based on analysis of data (usually interim data) from subjects in the study.

CBER, CDER FDA 2010

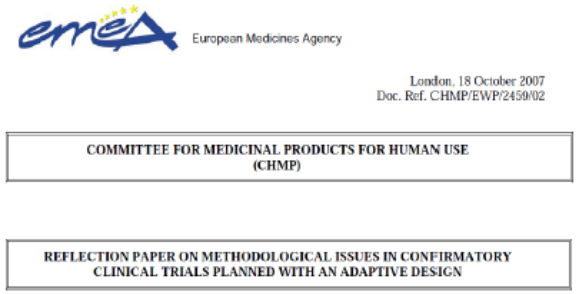

Draft Guidance for Industry and Food and Drug Administration Staff

Adaptive Designs for Medical Device Clinical Studies

A clinical study design that allows for prospectively planned modifications based on accumulating study data without undermining the trial's integrity and validity.

CBER, CDRH, FDA, 2016

Minimum requirements of adaptive designs


European Medicines Agency

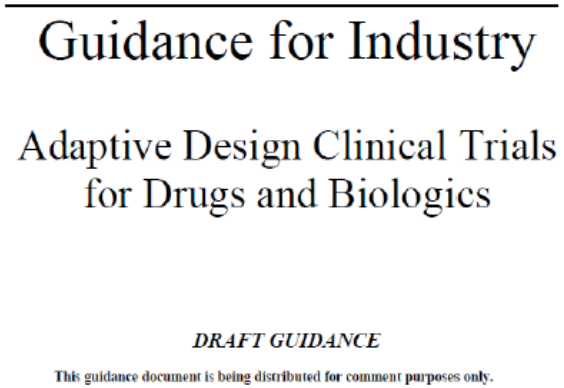
London, 18 October 2007
Doc. Ref. CHMP/EWP/2459/02

COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP)

REFLECTION PAPER ON METHODOLOGICAL ISSUES IN CONFIRMATORY CLINICAL TRIALS PLANNED WITH AN ADAPTIVE DESIGN

"Using an adaptive design implies that the statistical methods control the pre-specified type I error, that correct estimates and confidence intervals for the treatment effect are available, and that methods for the assessment of homogeneity of results from different stages are pre-planned."

EMA REFLECTION PAPER (2007)


Guidance for Industry

Adaptive Design Clinical Trials for Drugs and Biologics

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

"The chief concerns with these designs are control of the study-wide Type I error rate, minimization of the impact of any adaptation-associated statistical (see section VII.B) or operational bias on the estimates of treatment effects, and the interpretability of trial results."

FDA DRAFT GUIDANCE (2010)

The Adaptive designs CONSORT Extension (ACE) statement: a checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design

Munyaradzi Dimairo,[1] Philip Pallmann,[2] James Wason,[3,4] Susan Todd,[5] Thomas Jaki,[6] Steven A Julious,[1] Adrian P Mander,[2,3] Christopher J Weir,[7] Franz Koenig,[8] Marc K Walton,[9] Jon P Nicholl,[1] Elizabeth Coates,[1] Katie Biggs,[1] Toshimitsu Hamasaki,[10] Michael A Proschan,[11] John A Scott,[12] Yuki Ando,[13] Daniel Hind,[1] Douglas G Altman,[14] on behalf of the ACE Consensus Group

Table 1 | Some types of adaptations used in randomised trials with examples

Trial adaptive feature or adaptation, motivation, and cited examples of use	Type of adaptive design (AD) and examples of underlying statistical methods
Changing the predetermined sample size in response to inaccurate assumptions of study design parameters to achieve the desired statistical power.[38-40]	Sample size re-estimation, re-assessment, or re-calculation (SSR) using aggregated interim data from all participants or interim data separated according to allocated treatment.[41-48]
Stopping the trial early for efficacy, futility, or safety when there is sufficient evidence.[49 50]	Group sequential design (GSD)[51 52]; information-based GSD[53]; futility assessment using stochastic curtailment.[54-56]
Evaluating multiple treatments in one trial allowing for early selection of promising treatments or dropping futile or unsafe treatments.[57-59] New treatments can also be added to an ongoing trial.[60]	Multi-arm multi-stage (MAMS), dose/treatment-selection, drop-the-loser, or pick-the-winner, or add arm.[23 61-70]
Changing the treatment allocation ratio to favour treatments indicating beneficial effects.[71 72]	Response-adaptive randomisation (RAR).[72-77]
Investigating multiple research objectives that are traditionally examined in distinct trial phases, in one trial under a single protocol.[78-80] For instance, addressing learning (selecting promising treatments for further testing) and confirmatory objectives in one trial.	Operationally or inferentially seamless AD.[67-69 81-83]
Adjusting the trial population or selecting patients with certain characteristics that are most likely to benefit from investigative treatments.[84-87] This may involve incorporating statistical information from or adapting on a biomarker.	Population or patient enrichment or biomarker AD.[88-92]
Changing the primary research hypotheses or objectives or primary endpoints.[82 93] For example, switching from non-inferiority to superiority.	Adaptive hypotheses.[62 94]
Switching the allocated treatment of patients to an alternative treatment influenced by ethical considerations, for instance, due to lack of benefit or safety issues.	Adaptive treatment-switching.[95 96]
Combination of at least two types of adaptations.[24 40 93 97-102]	Multiple ADs such as GSD or drop-the-loser with SSR[103]; inferentially seamless phase 2/3 AD with hypotheses selection[81] or population enrichment[104]; biomarker-stratified with RAR[105]; adaptive platform trials where arms can be added or stopped early.[19 24 106]

thebmj

The Adaptive designs CONSORT Extension (ACE) statement: a checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design

Munyaradzi Dimairo,[1] Philip Pallmann,[2] James Wason,[3,4] Susan Todd,[5] Thomas Jaki,[6] Steven A Julious,[1] Adrian P Mander,[2,3] Christopher J Weir,[7] Franz Koenig,[8] Marc K Walton,[9] Jon P Nicholl,[1] Elizabeth Coates,[1] Katie Biggs,[1] Toshimitsu Hamasaki,[10] Michael A Proschan,[11] John A Scott,[12] Yuki Ando,[13] Daniel Hind,[1] Douglas G Altman,[14] on behalf of the ACE Consensus Group

For numbered affiliations see end of the article.
Correspondence to: M Dimairo m.dimairo@sheffield.ac.uk; mdimairo@gmail.com (ORCID 0000-0002-9311-6920)
Additional material is published online only. To view please visit the journal online.
Cite this as: *BMJ* 2020;369:m115 http://dx.doi.org/10.1136/bmj.m115
Accepted: 19 November 2019

Adaptive designs (ADs) allow pre-planned changes to an ongoing trial without compromising the validity of conclusions and it is essential to distinguish pre-planned from unplanned changes that may also occur. The reporting of ADs in randomised trials is inconsistent and needs improving. Incompletely reported AD randomised trials are difficult to reproduce and are hard to interpret and synthesise. This consequently hampers their ability to inform practice as well as future research and contributes to research waste. Better transparency and adequate reporting will enable the potential benefits of ADs to be realised.

This extension to the Consolidated Standards Of Reporting Trials (CONSORT) 2010 statement was developed to enhance the reporting of randomised AD clinical trials. We developed an Adaptive designs CONSORT Extension (ACE) guideline through a two-stage Delphi process with input from multidisciplinary key stakeholders in clinical trials research in the public and private sectors from 21 countries, followed by a consensus meeting. Members of the CONSORT Group were involved during the development process.

The paper presents the ACE checklists for AD randomised trial reports and abstracts, as well as an explanation with examples to aid the application of the guideline. The ACE checklist comprises seven new items, nine modified items, six unchanged items for which additional explanatory text clarifies further considerations for ADs, and 20 unchanged items not requiring further explanatory text. The ACE abstract checklist has one new item, one modified item, one unchanged item with additional explanatory text for ADs, and 15 unchanged items not requiring further explanatory text.

The intention is to enhance transparency and improve reporting of AD randomised trials to improve the interpretability of their results and reproducibility of their methods, results and inference. We also hope indirectly to facilitate the much-needed knowledge transfer of innovative trial designs to maximise their potential benefits.

"To maximise the benefit to society, you need to not just do research but do it well" *Douglas G Altman*

Purpose of the paper

Incomplete and poor reporting of randomised clinical trials makes trial findings difficult to interpret due to study methods, results, and inference that are not reproducible. This severely undermines the value of scientific research, obstructs robust evidence synthesis to inform practice and future research, and contributes to research waste.[1 2] The Consolidated Standards Of Reporting Trials (CONSORT) statement is a consensus-based reporting guidance framework that aims to promote and enhance transparent and adequate reporting of randomised trials.[3 4] Specific CONSORT extensions addressing the reporting needs for particular trial designs, hypotheses, and interventions have been developed.[5] The use of reporting guidelines is associated with improved completeness in study reporting[6-8]; however, mechanisms to improve adherence to reporting guidelines are needed.[9-12]

We developed an Adaptive designs CONSORT Extension (ACE)[13] to the CONSORT 2010 statement[14] to support reporting of randomised trials that use an adaptive design (AD)—referred to as AD randomised trials. In this paper, we define an AD and summarise some types of ADs as well as their use and reporting. We then describe briefly how the ACE guideline was developed, and present its scope and underlying principles. Finally, we present the ACE checklist with explanation and elaboration (E&E) to guide its use.

Adaptive designs: definition, current use, and reporting

The ACE Steering Committee[13] agreed a definition of an AD (box 1) consistent with the literature.[14-18]

Substantial uncertainties often exist when designing trials around aspects such as the target population, outcome variability, optimal treatments for testing, treatment duration, treatment intensity, outcomes to measure, and measures of treatment effect.[19] Well designed and conducted AD trials allow researchers to

Adaptive Two Stage Designs

Adaptive design allowing for design modification in one interim analysis

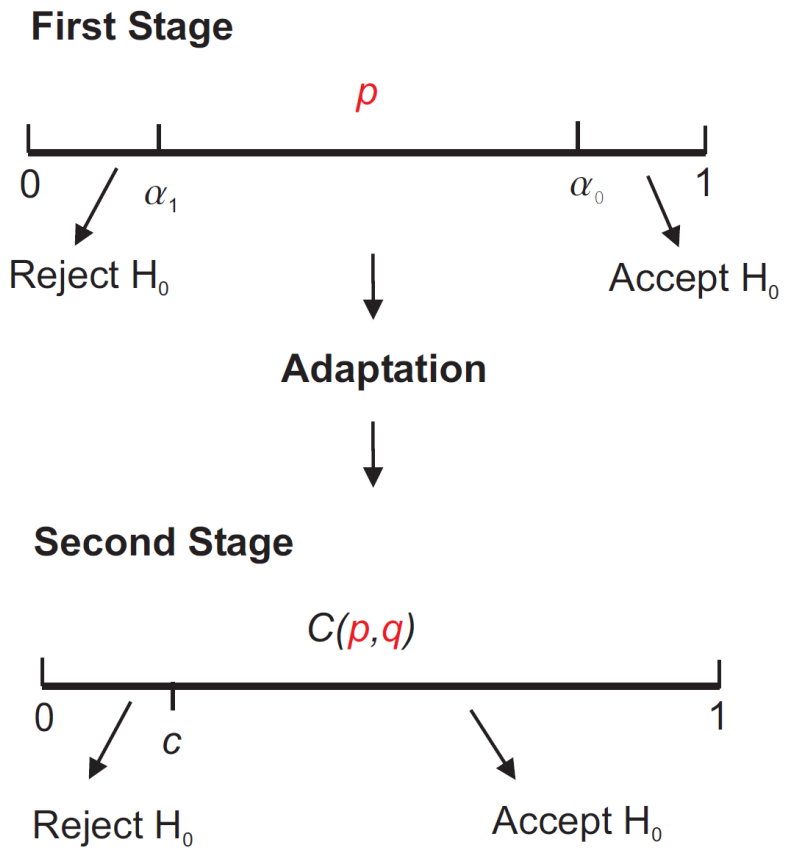
Adaptive Test Procedures

- A trial is performed in two stages
- In an interim analysis the trial may be
  - stopped for futility or efficacy or
  - continued and possibly adapted (sample size, test statistics)
- Adaptation of the design of second stage
  - adaptations depend on all (unblinded) interim data including secondary and safety endpoints.
  - the adaptation rule is not (completely) preplanned.

How to construct a test that controls the type I error?

- Tests Based on Combination Tests
- Tests Based on the Conditional Error Rate

Adaptive Combination Tests (Bauer 89, Bauer & Köhne 94)



Planning:

- Fix design (only) for Stage 1
- Fix combination function $C(p, q)$ and critical value c e.g. $C(p, q) = p \cdot q$

Stage 1:

- Compute p-value p from Stage 1 data
- Fix design for Stage 2 based on data from Stage 1

Stage 2:

- Compute p-value q form Stage 2 data.
- Reject H_0 iff $C(p, q) \leq c$.

Type I error control and combination functions

Type I error control

Type I error rate $\leq \alpha$ if we choose critical value c such that

$$P[p \leq \alpha or C(p, q) \leq c] = \alpha$$

for independent and uniformly distributed p-values p and q.

- *Fisher product test:* $C(p, q) = p \cdot q$
  (Bauer 1989, Bauer & Köhne, 1994)
- *Weighted inverse normal method:*
  $C(p, q) = \Phi(w_1 \Phi^{-1}(p) + w_2 \Phi^{-1}(q))$
  (Lehmacher & Wassmer, 1999)

(Remark: Can use critical values of a group sequential trial with interim information fraction w_1).

Clue of Adaptive Tests

- Do not pool the data of the stages, combine the stage-wise p-values.
- Then the distribution of the combination function under the null does not depend on design modifications
- Hence the adaptive test is still a test at the level alpha for the modied design!
- Applicable also for multiple looks, multiple hypotheses, ...
- Adaptations can depend on all (unblinded) interim data including secondary and safety endpoints.
- For a control of the type I error rate, one need not pre-specify how the Stage 1 data determine the design of Stage 2.

Numerical Example Product Test

One sample test at level $\alpha = 0.025$ for the mean of (pre-planned) 40 normally distributed observations to test the hypotheses

$$H_0 : \mu = 0 \quad \text{against} \quad H' : \mu > 0$$

- Product test $\alpha_1 = 0.01, \alpha_0 = 1, c = 0.00326$.
- First stage sample size $n_1 = 20$ observations.
- First stage data: mean 3.7, sd 10.9, $p = 0.0727$ (t-test).
- Interim decision: $p > \alpha_1$ continue.
- Second stage: Choose sample size of $n^{(2)} = 30$ observations.
- Second stage data: mean 3.2, sd 9.5, $q = 0.0376$ (t-test).
- Test decision: $p \cdot q = 0.00273 < c$ reject H_0.

Multiplicity in Adaptive Clinical Trials

Multiplicity arises through

- multiple treatment groups
- multiple endpoints
- multiple subgroups

In Adaptive Clinical Trials, treatment groups, endpoints and subgroups may be dropped or added in interim analyses while controlling the Familywise Type I Error Rate in the strong sense.

Control of the familywise type I error rate in the strong sense:

The probability that any true null hypothesis is rejected is bounded by α, regardles of which and how many null hypotheses are true.

Adaptive Seamless Designs

Adaptive Designs with treatment selection at an interim analysis

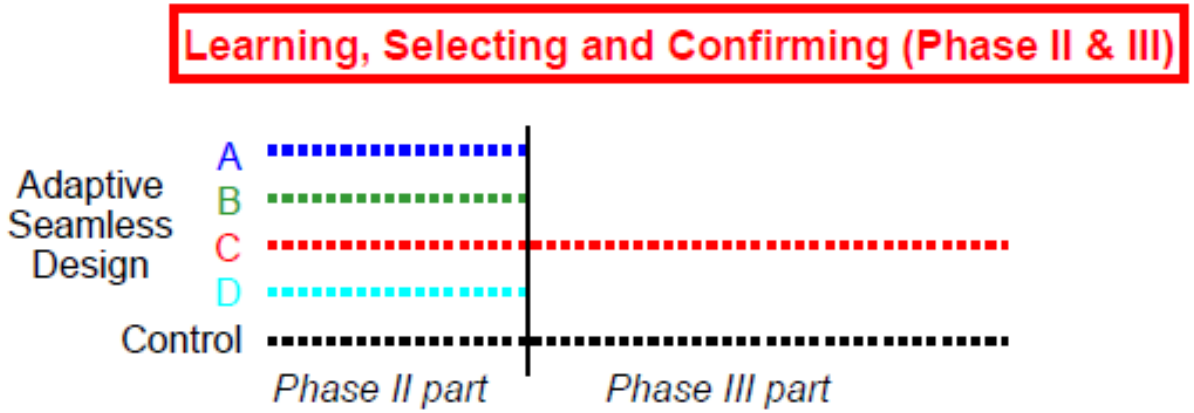
Phases of Clinical Development

- PHASE I TRIALS determine biological mechanism and side effects in humans (probands or patients).
- PHASE II TRIALS explore therapeutic effect and short-term side effects; learning for phase III, e.g., determine dose(s) used in Phase III trial.
- PHASE III TRIALS shall confirm preliminary evidence for effectiveness and safety of the drug and provide adequate information for marketing approval → hypothesis testing.
- PHASE IV TRIALS: Post-marketing studies to delineate additional information on drug's risks, benefits, and use.

Separate Phase II and III Trials

- Conduct phase II trial.
- Plan phase III trial based on the information from phase II trial (which treatment, which number of patients, etc.).
- Conduct confirmatory phase III trial. Demonstrate efficacy using ONLY phase III trial data.

Adaptive Seamless Phase II + III Trials



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

Adaptive Seamless Phase II + III Trials

- Smaller time lag between phase II and phase III. Speeds up the drug development process.
- Allows us to use also the data form a (late) Phase II trial for the efficacy hypothesis testing. This saves resources (patients), costs and time.
- Improves quality of the drug development process by using the same study protocol (study plan) for the (late) phase II and (early) phase III trial.

Case Study I: Interim Dose Selection (see Elsässer 2014)

Proposal for study seeking scientific advice (SAWP) from European Medicine Agency (EMA)

- Seamless phase II/III designs for two pivotal placebo controlled trials of a new chemical entity for the treatment of diabetic nephropathy
- Objectives:
  - Demonstrate superiority in a surrogate marker of kidney disease progression
  - Select two of three initially tested dose strengths based on an interim analysis of the benefit/risk ratio in both trials.
- Pre-planned interim analyses to be performed by an IDMC after 60% of 420 patients had completed 8 weeks of treatment in the first trial.
- Dose selection based on data from both trials using pre-determined criteria for the primary efficacy and safety parameters.
- Proposed type I error rate control: Bonferroni adjustment to control the familywise error rate adjusting the level for two comparisons only.

Case Study I: EMA (SAWP/CHMP) Reply

- The statistical testing procedure was not endorsed, as it was not supposed to control the familywise type I error rate for the three hypotheses initially considered.
- Instead, adaptive combination tests based on the closure principle and adaptive Dunnett test procedures based on the conditional error rate are adequate methods to control the type I error rate.
- The advantage of the proposed design with respect to power should be evaluated as it maybe small.
- Safety evaluation may not be possible to support dose selection at the proposed time of interim analysis.

Case Study I: Issues to address

- Control of the type I error rate *Even if only one experimental arm (and the control) is selected, a multiplicity correction is required.*

  EMA Reflection Paper, 2007, FDA Draft Guidance, 2010

- Interim selection of treatments may introduce bias
  If the treatment with the largest interim effect size is chosen, the effect estimates will be biased.
- Interim data maybe highly variable and lead to selection of the "wrong" treatment arm.
- Unblinding of data at the interim analysis

Case Study I: Control of Type I error rate

Several procedures have been proposed:

- Methods based on completely predefined adaptation rules
  - Multiplicity adjusted critical values are determined by simulation or numerical integration

Thall et al. '88, '89, Stallard and Todd '03, Sampson and Sill '05, Magirr et al. '12 ...

- Combine Closure Principle and Adaptive Designs
  - Perform adaptive tests for intersection hypotheses using

Bauer and Kieser '99, Kieser et al. '99, Hommel 2001, Posch et al. '05, König et al. '08, Bretz et al. '09, Posch et al. '11

Adaptive Designs based on the closure principle

- Selection of treatments may depend on all data collected (also safety data, secondary endpoints)
- Sample sizes may be adapted.
- In principle, pre-specification of the adaptation rules is not required to control the multiple Type I error rate.
  - However, the type of adaptations and the anticipated adaptation rules should be pre-specified
  - Number of adaptations should be limited

Dose selection and efficacy testing

- Parallel group design with $k = 2$ dose groups and a control group (i.e., in total three parallel groups).
- Testing the one sided hypotheses

Dose 1 vs control: $H_{0,1} : \mu_1 \leq \mu_0$ vs. $H_{1,1} : \mu_1 > \mu_0$

Dose 2 vs control: $H_{0,2} : \mu_2 \leq \mu_0$ vs. $H_{1,2} : \mu_2 > \mu_0$

Dose selection and efficacy testing

- After Stage 1 we decide either to
  - go into Stage 2 with BOTH doses or
  - go into Stage 2 with only ONE dose.
- Selection rule unknown before end of Stage 1.
- Choice of sample sizes for Stage 2 depends on selected dose(s) and observed efficiency.
- Regulatory bodies ask for a level $\alpha = 0.025$ test of the intersection hypothesis

$$H_{0,1} \cap H_{0,2} : \mu_1, \mu_2 \leq \mu_0$$

Flexible Closed Test (Bauer & Kieser 1999, Hommel 01)

- Use flexible two stage test for $H_{0,1} \cap H_{0,2}$,
  e.g. fix a combination test $C(p, q)$ at level α.
- At Stage 1 use a multiplicity adjusted p-value for p
  e.g. p-value of Šidak test

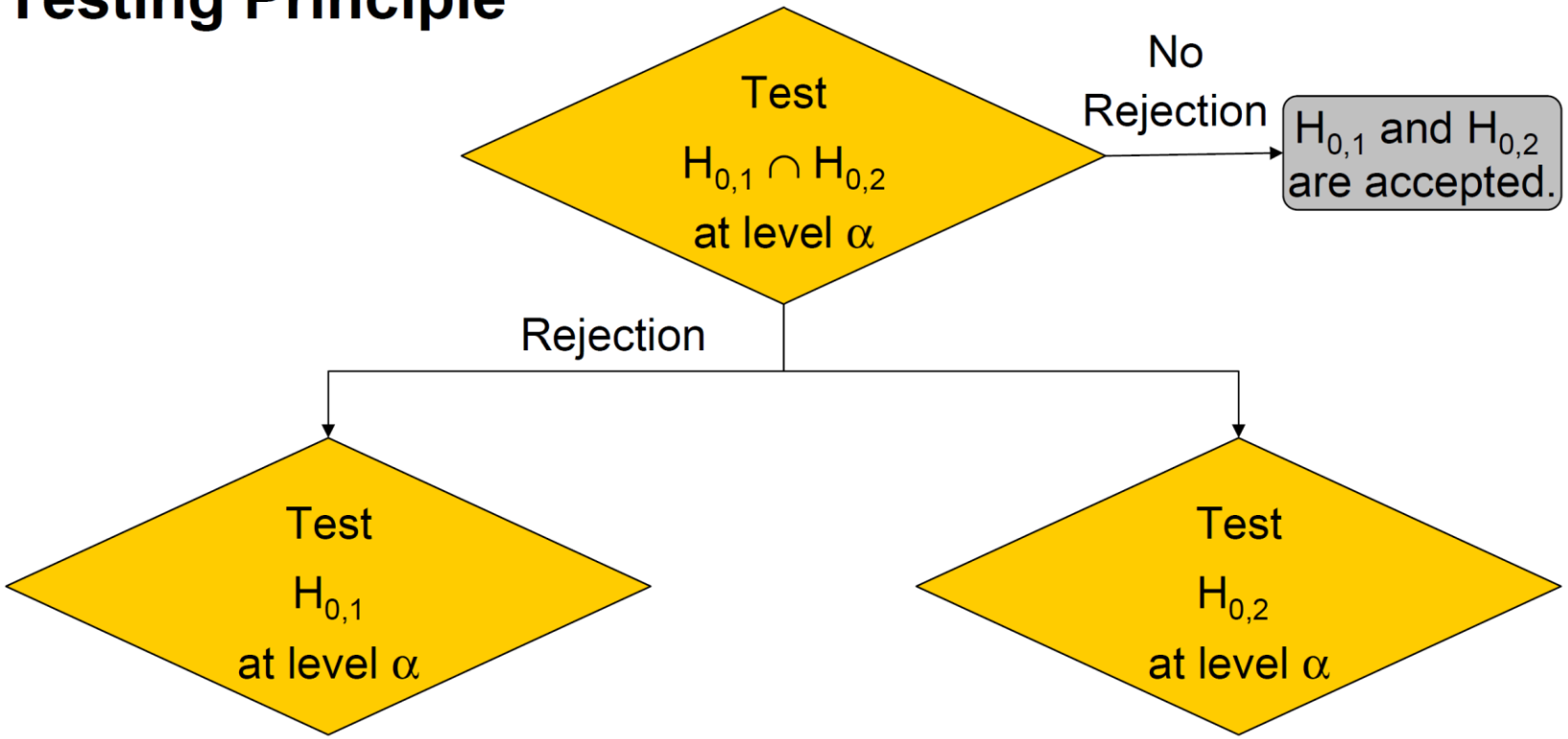
$$p = p_{12} = 1 - [1 - \min(p_1, p_2)]^2$$

- At Stage 2 use the p-value for the selected doses(s):
  – If we select only one, e.g. dose 1, we use $q = q_1$
  – If we select both, we use e.g. Šidak test

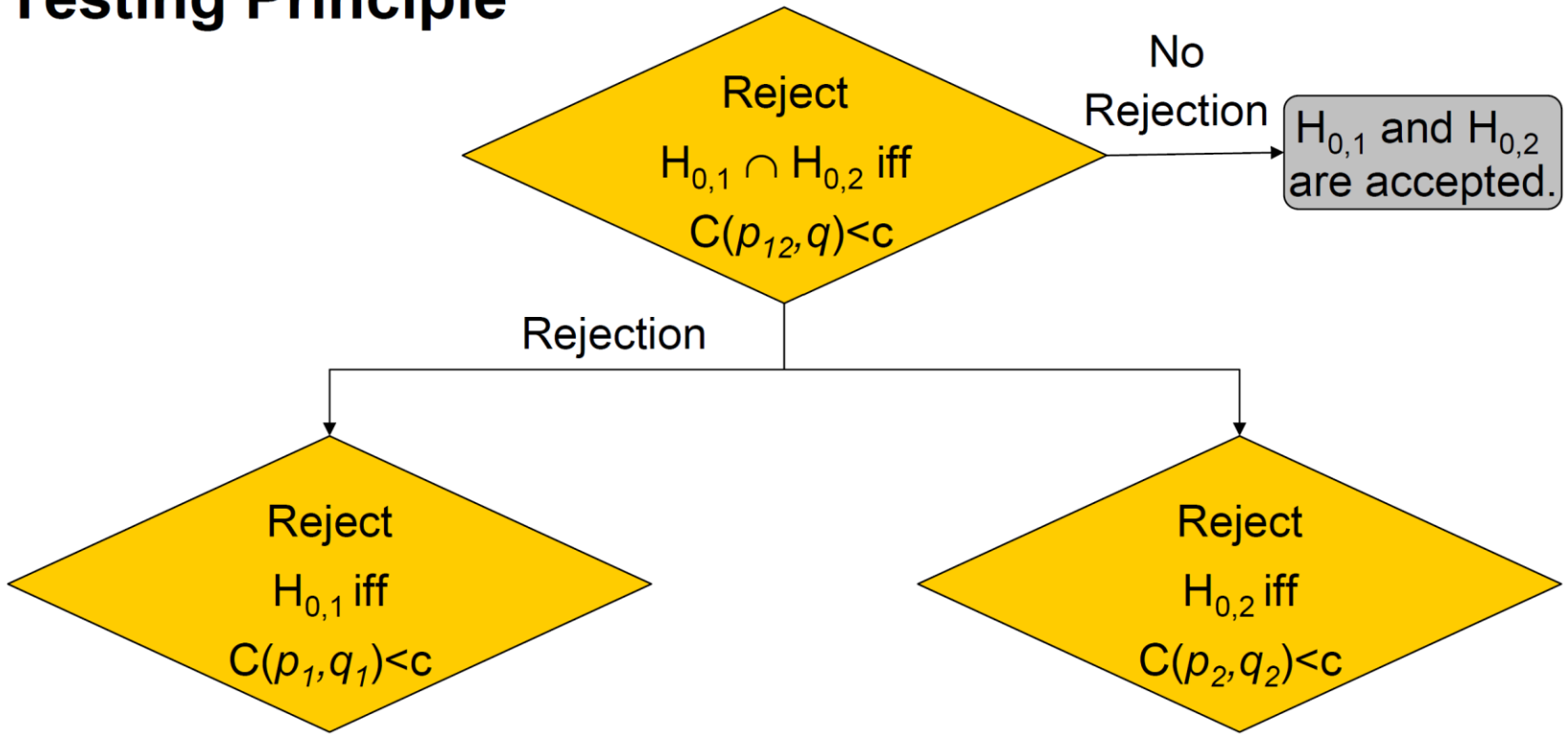
$$q = q_{12} = 1 - [1 - \min(q_1, q_2)]^2$$

- In all cases reject $H_{0,1} \cap H_{0,2}$ iff $C(p, q) \leq c$.

The Closed Testing Principle



Adaptive Closed Testing Principle

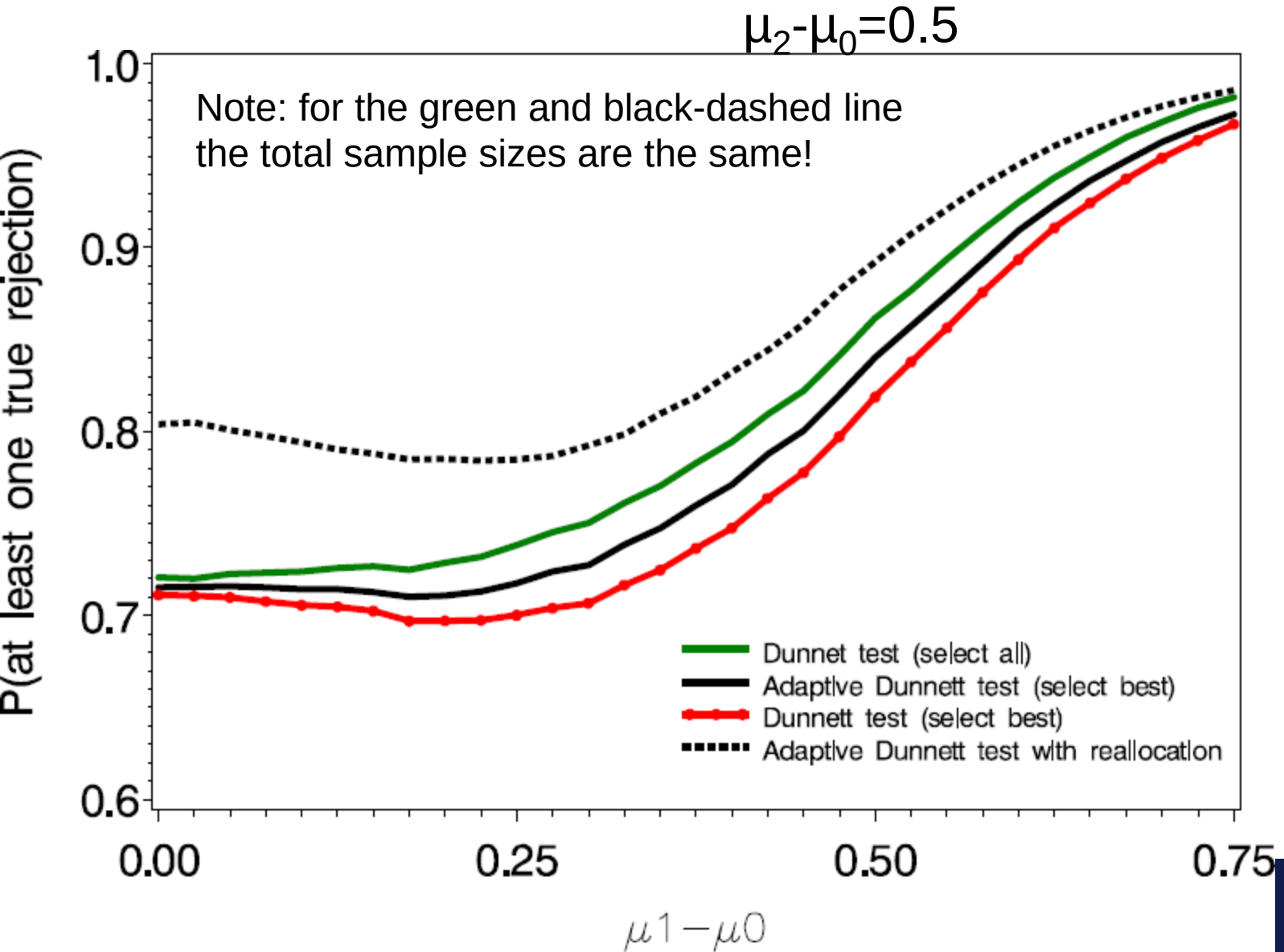


Power considerations

How to define the "power" if multiple hypotheses are tested in an adaptive trial?

- Probability to reject all hypotheses.
- Probability to reject all selected hypotheses.
- Average power for the selected hypotheses.
- The probability for a particular treatment to be selected and the corresponding hypothesis to be rejected.
- Probability to select and reject any hypotheses.

Selecting the treatment with larger effect at Stage One and sample size reallocation



- Two treatments versus control
- Normal responses ($\sigma = 1$)
- Total n such that power for single treatment-control comparison is 80% for $\mu_i - \mu_0 = 0.5$
- Interim analysis at $n_1 = n/2$ with selection of "best" treatment
- mean diff. for treatment 2:

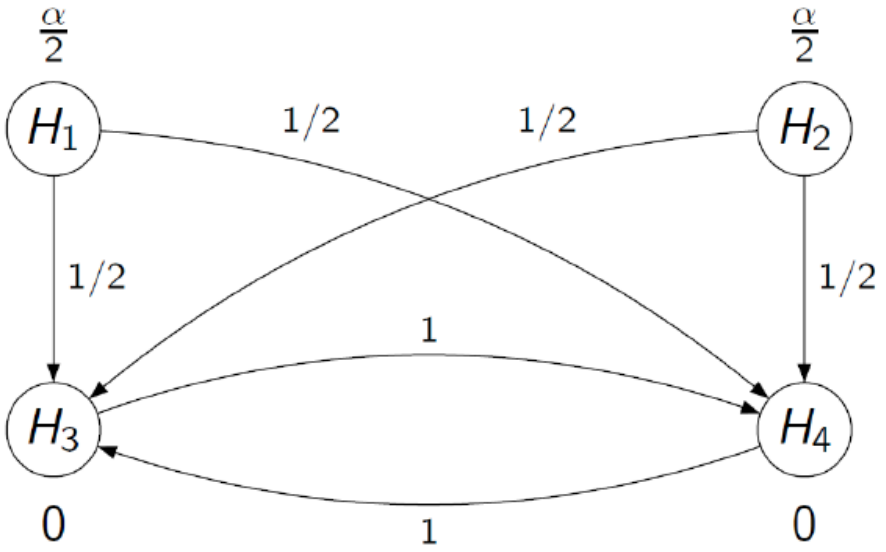
$$\mu_2 - \mu_0 = 0.5$$

Koenig et al. 2008, Koenig et al. 2006, Klinglmüller, Posch & *Koenig F* (2014), Hlavin, Hampson, Koenig (2017), Bretz & *Koenig* et al. (2009), …

What if the tests become too complex?

- E.g., 2 Primary and 2 Secondary Hypotheses
- Initially, test the primary hypothesis
- If at least one primary hypothesis can be rejected, continue to test the secondary hypotheses
- With which (adaptive) test can the secondary hypotheses be tested?
- Complex closed adaptive test with many (intersection) hypothes

$H_1 \cap H_2 \cap H_3 \cap H_4$
$H_1 \cap H_2 \cap H_3$
$H_1 \cap H_2 \cap H_4$
$H_1 \cap H_2$
$H_1 \cap H_3 \cap H_4$
$H_1 \cap H_3$
$H_1 \cap H_4$
H_1
$H_2 \cap H_3 \cap H_4$
$H_2 \cap H_3$
$H_2 \cap H_4$
H_2
$H_3 \cap H_4$
H_3
H_4



- Fixed Sample: Bretz, Maurer, Brannath & Posch Statistics in Medicine (2009), Bretz, Posch, Glimm, Klinglmueller, Maurer, & Rohmeyer, K. (2011). Biometrical Journal
- Adaptive graph-based multiple testing procedures. Klinglmueller, Posch, Koenig (2014) DOI: 10.1002/pst.1640

- Use graphical methods to define and derive appropriate adaptive tests!

Further application: Adaptive Enrichment Designs

- Instead of selecting treatment arms, subgroups can be dropped or added at an adaptive interim analysis

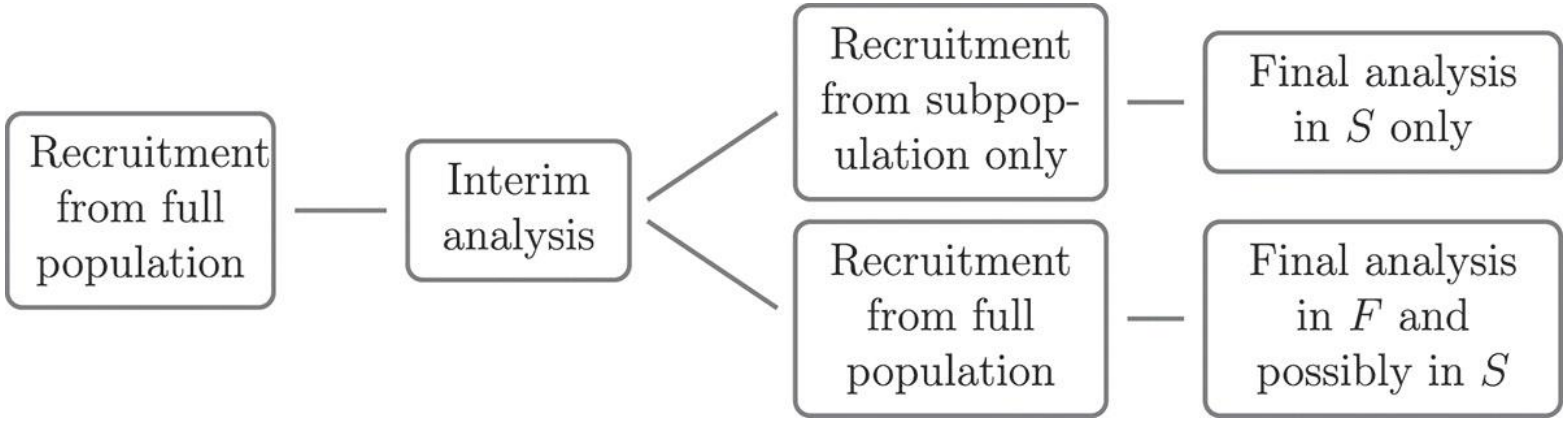


Figure from Ondra et al. 2018 https://www.tandfonline.com/doi/full/10.1080/10543406.2015.1092034

Ballarini, N. M., Burnett, T., Jaki, T., Jennison, C., König, F., & Posch, M. (2021). Optimizing subgroup selection in two-stage adaptive enrichment and umbrella designs. *Statistics in medicine*, *40*(12), 2939-2956.

Graf, A. C., Posch, M., & Koenig, F. (2015). Adaptive designs for subpopulation analysis optimizing utility functions. *Biometrical Journal*, *57*(1), 76-89.

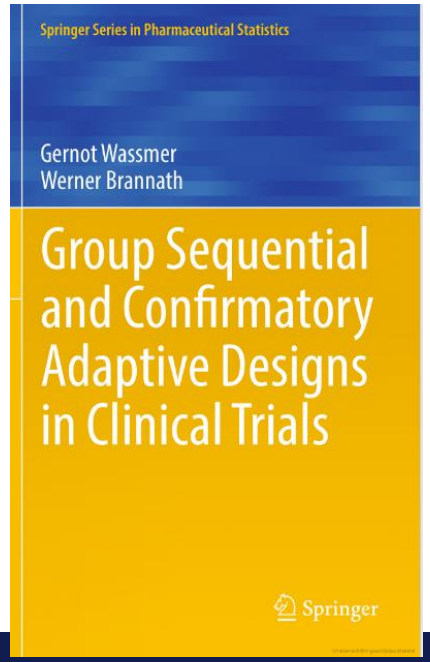
Ondra, T., Jobjörnsson, S., Beckman, R. A., Burman, C. F., König, F., Stallard, N., & Posch, M. (2019). Optimized adaptive enrichment designs. *Statistical methods in medical research*, *28*(7), 2096-2111.

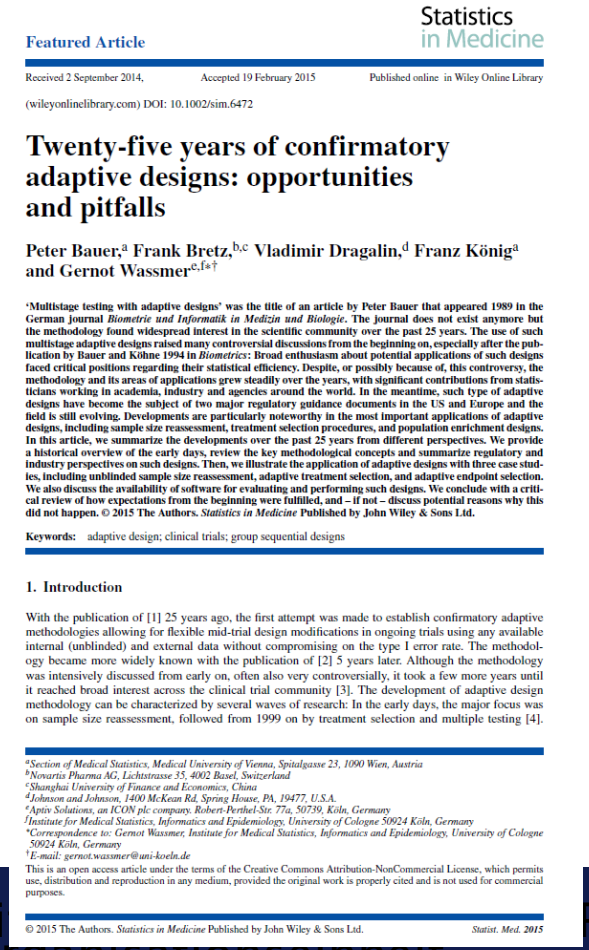
Ballarini, Nicolás M., et al. "A critical review of graphics for subgroup analyses in clinical trials." *Pharmaceutical Statistics* 19.5 (2020): 541-560.

Sugitani, T., Posch, M., Bretz, F., & Koenig, F. (2018). Flexible alpha allocation strategies for confirmatory adaptive enrichment clinical trials with a prespecified subgroup. *Statistics in Medicine*, *37*(24), 3387-3402.

Some references on adaptive designs

- P. Bauer, F. Bretz, V. Dragalin, F. Koenig, and G. Wassmer
  Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls
  Statistics in Medicine **35**, 325-347, (2016)
  http://dx.doi.org/10.1002/sim.6472
- F. Bretz, F. Koenig, W. Brannath, E. Glimm, and M. Posch
  Adaptive designs for confirmatory clinical trials
  Statistics in Medicine **28**, 1181--1217, 2009
  http://dx.doi.org/10.1002/sim.3538





Featured Article — Statistics in Medicine

Received 2 September 2014, Accepted 19 February 2015 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6472

Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls

Peter Bauer,[a] Frank Bretz,[b,c] Vladimir Dragalin,[d] Franz König[a] and Gernot Wassmer[e,f*†]

'Multistage testing with adaptive designs' was the title of an article by Peter Bauer that appeared 1989 in the German journal *Biometrie und Informatik in Medizin und Biologie*. The journal does not exist anymore but the methodology found widespread interest in the scientific community over the past 25 years. The use of such multistage adaptive designs raised many controversial discussions from the beginning on, especially after the publication by Bauer and Köhne 1994 in *Biometrics*: Broad enthusiasm about potential applications of such designs faced critical positions regarding their statistical efficiency. Despite, or possibly because of, this controversy, the methodology and its areas of applications grew steadily over the years, with significant contributions from statisticians working in academia, industry and agencies around the world. In the meantime, such type of adaptive designs have become the subject of two major regulatory guidance documents in the US and Europe and the field is still evolving. Developments are particularly noteworthy in the most important applications of adaptive designs, including sample size reassessment, treatment selection procedures, and population enrichment designs. In this article, we summarize the developments over the past 25 years from different perspectives. We provide a historical overview of the early days, review the key methodological concepts and summarize regulatory and industry perspectives on such designs. Then, we illustrate the application of adaptive designs with three case studies, including unblinded sample size reassessment, adaptive treatment selection, and adaptive endpoint selection. We also discuss the availability of software for evaluating and performing such designs. We conclude with a critical review of how expectations from the beginning were fulfilled, and – if not – discuss potential reasons why this did not happen. © 2015 The Authors. *Statistics in Medicine* Published by John Wiley & Sons Ltd.

Keywords: adaptive design; clinical trials; group sequential designs

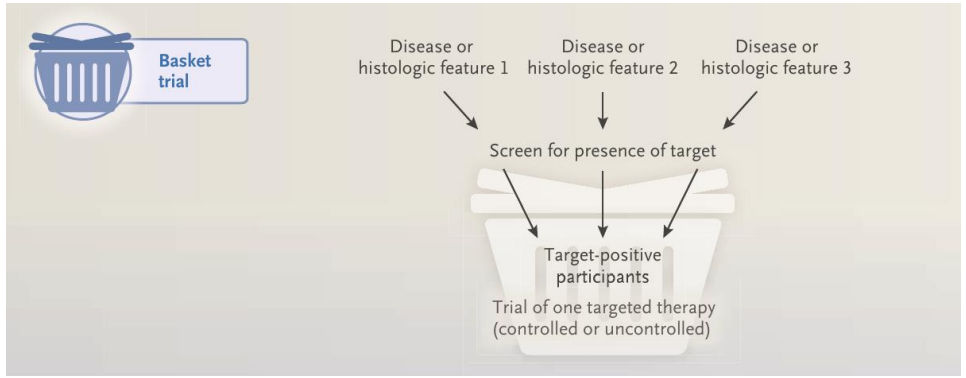
1. Introduction

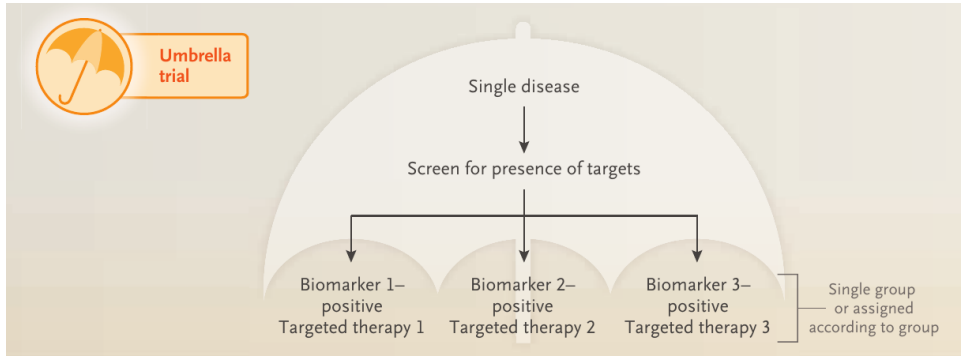
With the publication of [1] 25 years ago, the first attempt was made to establish confirmatory adaptive methodologies allowing for flexible mid-trial design modifications in ongoing trials using any available internal (unblinded) and external data without compromising on the type I error rate. The methodology became more widely known with the publication of [2] 5 years later. Although the methodology was intensively discussed from early on, often also very controversially, it took a few more years until it reached broad interest across the clinical trial community [3]. The development of adaptive design methodology can be characterized by several waves of research: In the early days, the major focus was on sample size reassessment, followed from 1999 on by treatment selection and multiple testing [4].

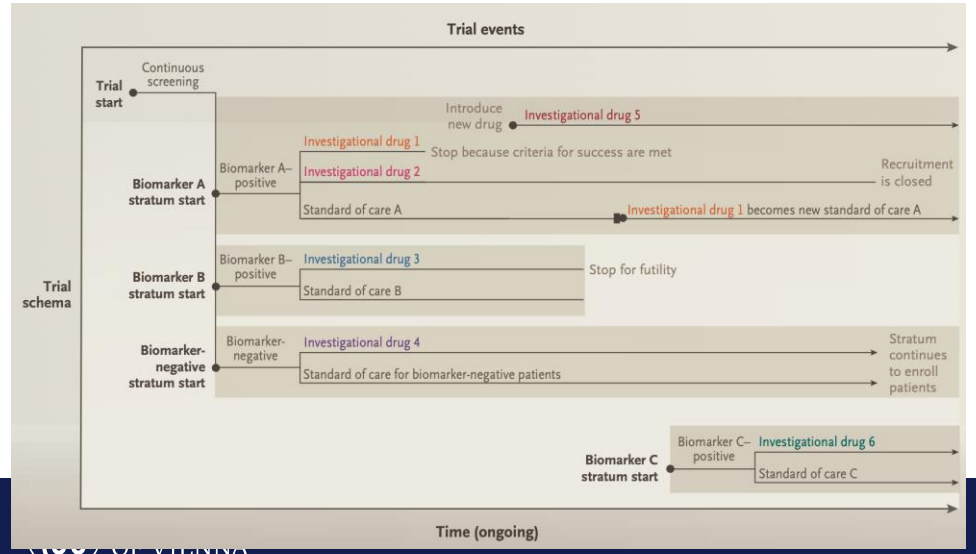
[a]*Section of Medical Statistics, Medical University of Vienna, Spitalgasse 23, 1090 Wien, Austria*
[b]*Novartis Pharma AG, Lichtstrasse 35, 4002 Basel, Switzerland*
[c]*Shanghai University of Finance and Economics, China*
[d]*Johnson and Johnson, 1400 McKean Rd. Spring House, PA, 19477, U.S.A.*
[e]*Aptiv Solutions, an ICON plc company. Robert-Perthel-Str. 77a, 50739, Köln, Germany*
[f]*Institute for Medical Statistics, Informatics and Epidemiology, University of Cologne 50924 Köln, Germany*
**Correspondence to: Gernot Wassmer, Institute for Medical Statistics, Informatics and Epidemiology, University of Cologne 50924 Köln, Germany*
†E-mail: gernot.wassmer@uni-koeln.de

Master Protocols

Master Protocols







- **Basket trial:** One investigational treatment (combination) is evaluated in the context of multiple diseases or disease subtypes with a common therapeutic target

- **Umbrella Trial**: Multiple investigational treatments (combinations) are evaluated in the context of a single disease, possibly within several substudies for different disease subtypes

- **Platform trial:** Umbrella trial, where drugs (combinations) may enter or leave the trial (e.g., if a new biomarker to identify disease subtypes becomes available)

Woodcock and LaVange ’17

Systematic Literature Review: https://doi.org/10.1016/j.clinthera.2020.05.010

Clinical Therapeutics/Volume 42, Number 7, 2020

The Evolution of Master Protocol Clinical Trial Designs: A Systematic Literature Review

Elias Laurin Meyer, MSc[1]; Peter Mesenbrink, PhD[2]; Cornelia Dunger-Baldauf, PhD[3]; Hans-Jürgen Fülle, MD PhD[3]; Ekkehard Glimm, PhD[3]; Yuhan Li, M.S.[2]; Martin Posch, PhD[1]; and Franz König, PhD[1]

[1]*Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria;* [2]*Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA;* and [3]*Novartis Pharma AG, Basel, Switzerland*

ABSTRACT

Purpose: Recent years have seen a change in the way that clinical trials are being conducted. There has been a rise of designs more flexible than traditional adaptive and group sequential trials which allow the investigation of multiple substudies with possibly different objectives, interventions, and subgroups conducted within an overall trial structure, summarized by the term *master protocol*. This review aims to identify existing master protocol studies and summarize their characteristics. The review also identifies articles relevant to the design of master protocol trials, such as proposed trial designs and related methods.

Methods: We conducted a comprehensive systematic search to review current literature on master protocol trials from a design and analysis perspective, focusing on platform trials and considering basket and umbrella trials. Articles were included regardless of statistical complexity and classified as reviews related to planned or conducted trials, trial designs, or statistical methods. The results of the literature search are reported, and some features of the identified articles are summarized.

Findings: Most of the trials using master protocols were designed as single-arm (n = 29/50), Phase II trials (n = 32/50) in oncology (n = 42/50) using a binary endpoint (n = 26/50) and frequentist decision rules (n = 37/50). We observed an exponential increase in publications in this domain during the last few years in both planned and conducted trials, as well as relevant methods, which we assume has not yet reached its peak. Although many operational and statistical challenges associated with such trials remain, the general consensus seems to be that master protocols provide potentially enormous advantages in efficiency and flexibility of clinical drug development.

Implications: Master protocol trials and especially platform trials have the potential to revolutionize clinical drug development if the methodologic and operational challenges can be overcome. (*Clin Ther.* 2020;42:1330–1360) © 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

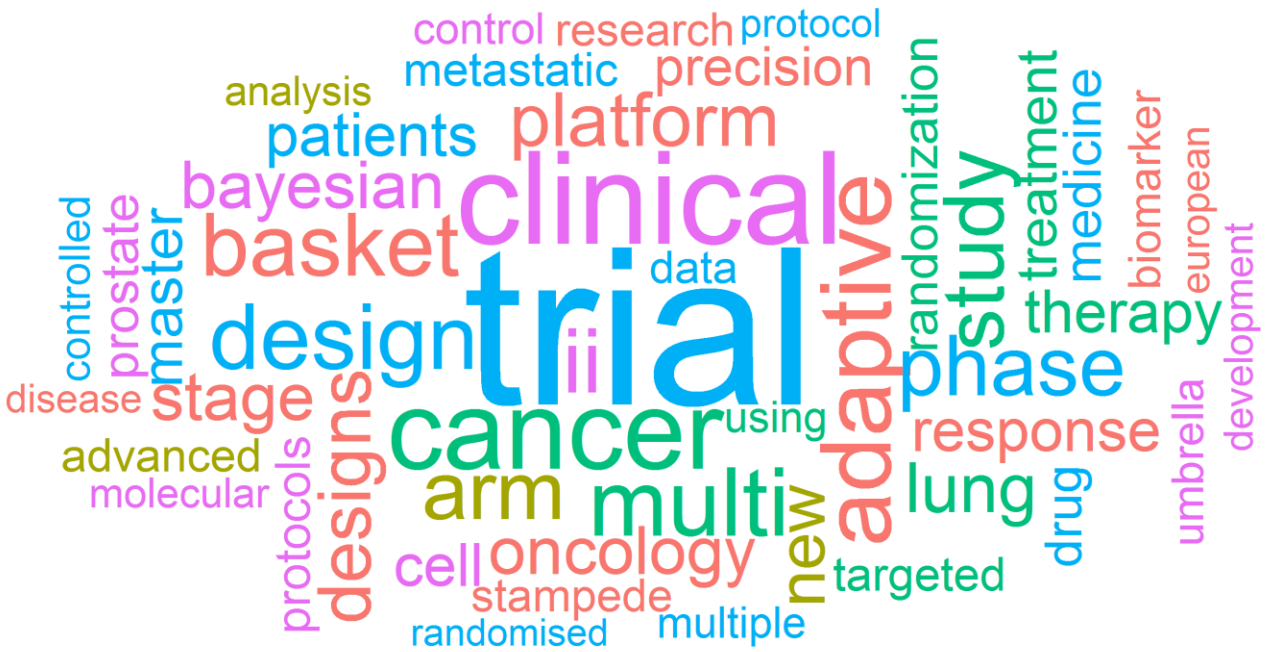
Key words: adaptive design basket trial, master protocol, multi-arm multi-stage design, platform trial.

INTRODUCTION

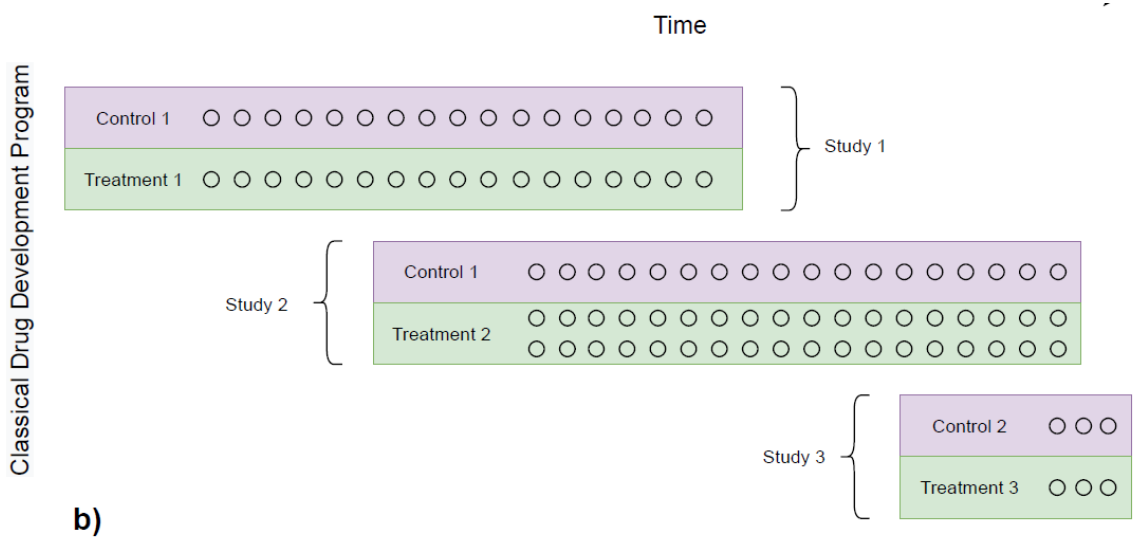
Since the late 1940s, randomized controlled trials (RCTs) have served as the gold standard for establishing therapeutic efficacy.[1] However, recent advances in drug discovery and biotechnology have accelerated tremendously the detection of treatment candidates. In addition, diagnostics have become more refined, leading to more precisely defined disease descriptions and hence smaller patient populations for targeted therapies. The classic 2-arm parallel-group RCTs have thus become one of the rate-limiting factors in drug development, and more

Accepted for publication May 11, 2020
https://doi.org/10.1016/j.clinthera.2020.05.010
0149-2918/$ - see front matter

- literature search on PubMed last updated January 01, 2020, search terms such
  - master protocol*[Title/Abstract] OR
  - platform/basket/umbrella trial/stud/design*[Title/Abstract] OR
- Included 164/678 identified papers + 122 manually
- In total 50 planned or conducted trials with master protocol identified



Classical Drug Development Programs



Traditionally:

- Type 1 error (T1E) control at study level
- No data-sharing accross studies
- Sample size / power calculations quite simple
- Don´t share information accross studies / indications etc

Why is there the wish for something different?

- Inefficient usage of resources
  - Standalone RCTs need their own control group
  - Each time develop new protocol, SAP,
  - Seek ethics & regulatory approval,
  - Look for appropriate trial sites, …
- Advances in personalized medicine lead to massive amount of hypotheses

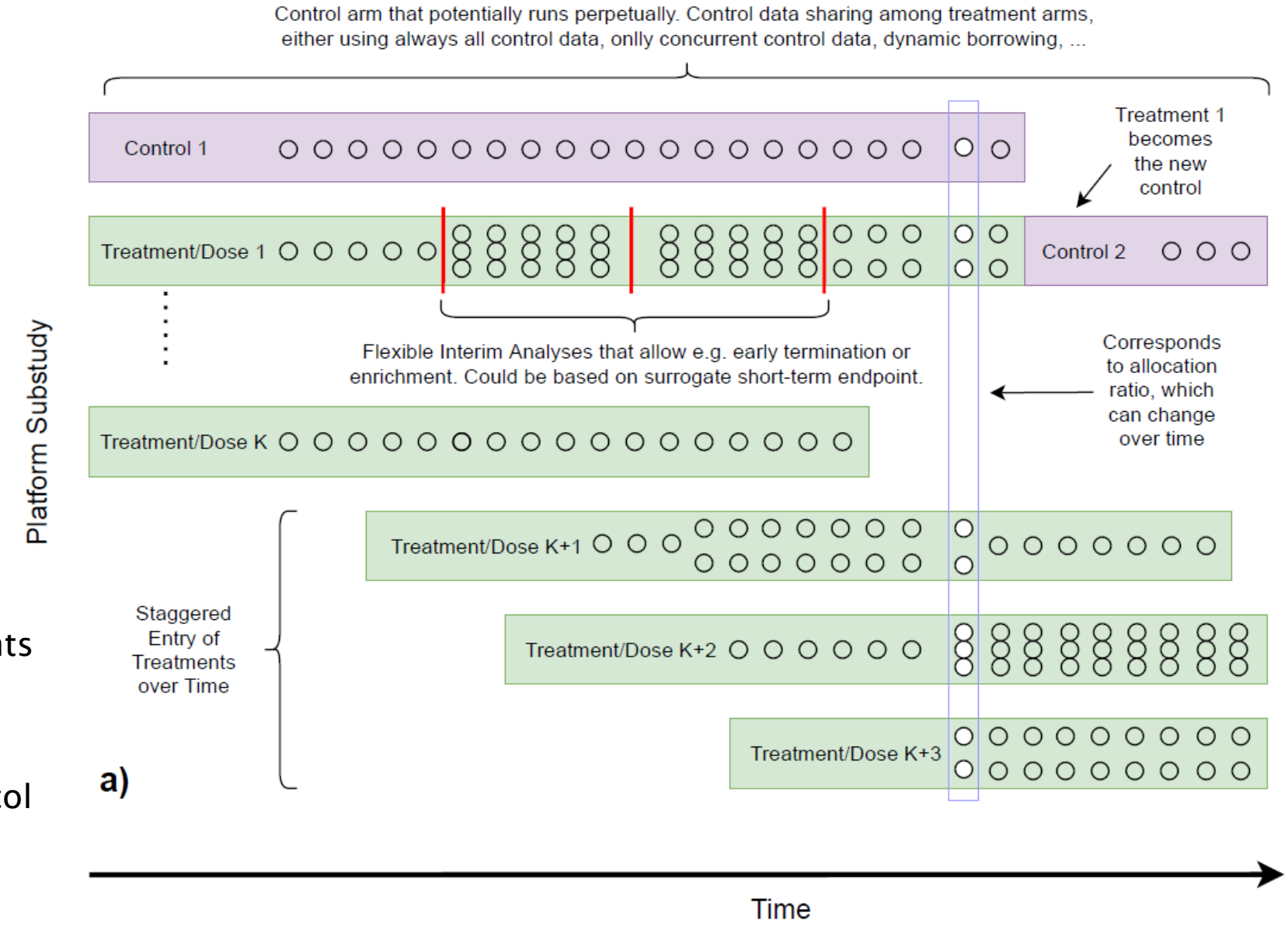
Koenig et al. (2024), Meyer et al. (2020b)

Collaborative Platform Trials

Koenig et al. (2024), Meyer et al. (2020b)

Design Characteristics of Platform Trials

- Multi-armed trials
- Interim analyses & adaptations
- Treatments to be studied not defined upfront but may enter during the course of the trial
- Control arm(s) can be shared
- Control arm(s) may change over time
- Populations for the different treatments may not be the same (Umbrella type trials)
- Designed as trial with a Master Protocol with several sub-studies



Potential advantages of platform trial

Operational:

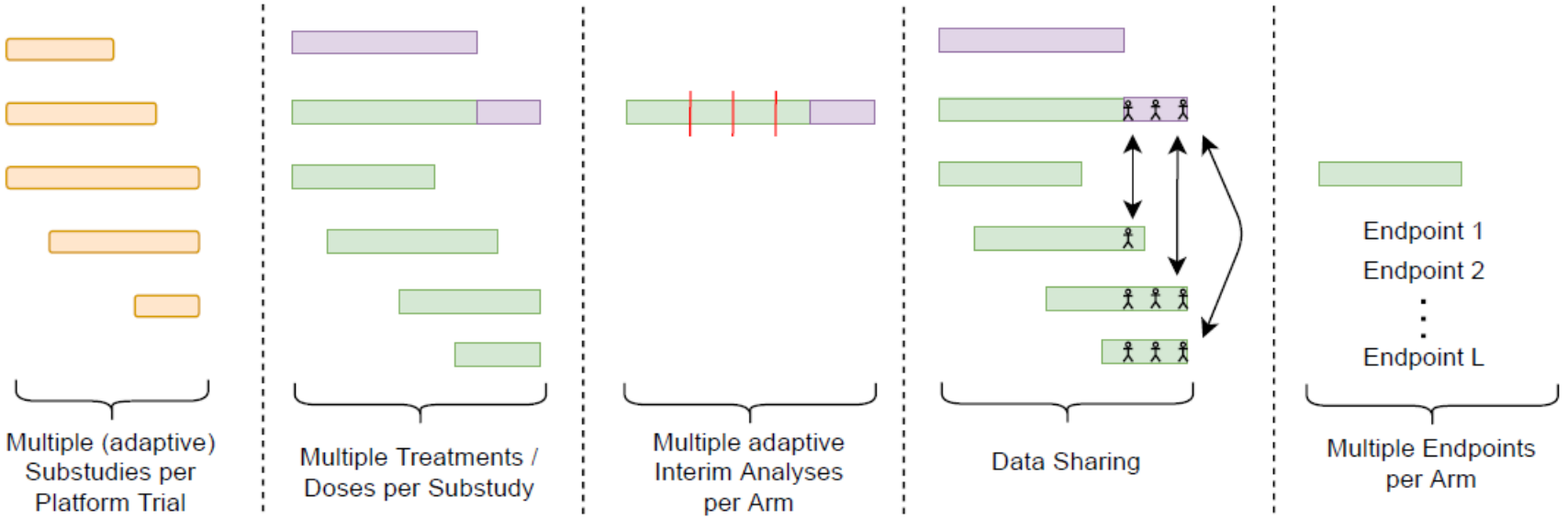
- More patients eligible for trial due to multiple treatments and sub-studies with possibly different inclusion criteria
- Joint trial infrastructure leads to savings in time and money for sponsor(s)

Statistical:

- Multiple hypotheses tested in the same trial (which is also a big challenge)
- Sharing of control data and adaptive decision rules potentially lead to fewer number of patients required
- Direct comparison between treatments allows for adaptive randomization leading to effective treatments “graduating” faster and fewer patients on inefficacious treatments

Multiplicity Issues in Platform Trials

Sources of structural multiplicty in Platform trials



Multi-regional different regulatory requirements

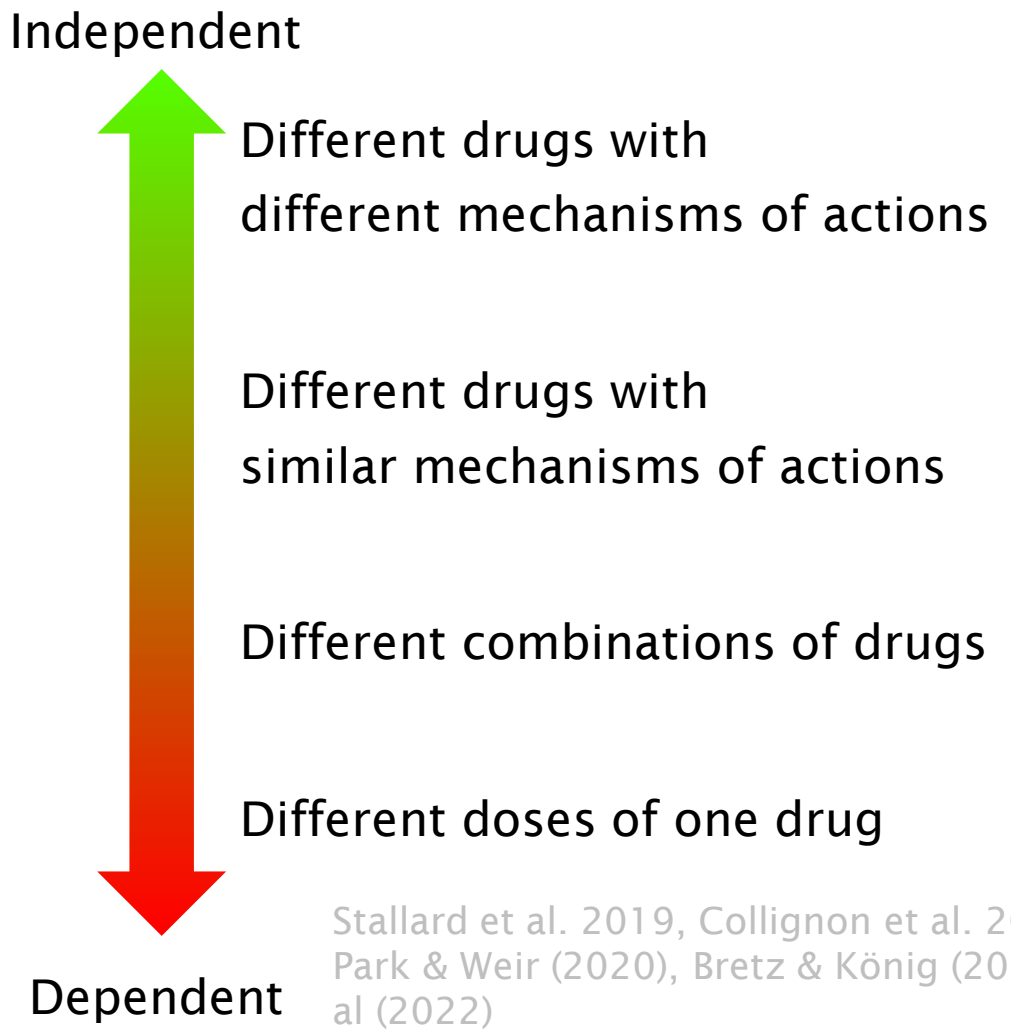
Multiple interventions + *Adaptive decision making* + *Multiple Control Groups* + *Multiple endpoints* + *Multiple subgroups* + *Multi-regional*

= *Large convoluted multiplicity problem*

Is there a need to adjust for multiplicity in Platform trials?

NO need to adjust WHEN hypotheses are **inferentially** independent

- Hypotheses are inferentially independent, if **the truth or falsehood of one hypothesis is unrelated to the truth and falsehood of the other hypotheses.**
- no extrapolation from one hypotheses to the the other is possible.
- If we did separate trials, we would also not adjust for multiplicity (and the shared control group leads to a lower FWER anyway)



Stallard et al. 2019, Collignon et al. 2020a, 2020b, Park & Weir (2020), Bretz & König (2020), Nguyen et al (2022)
EU-PEARL session on multiplicity first stakeholder workshop

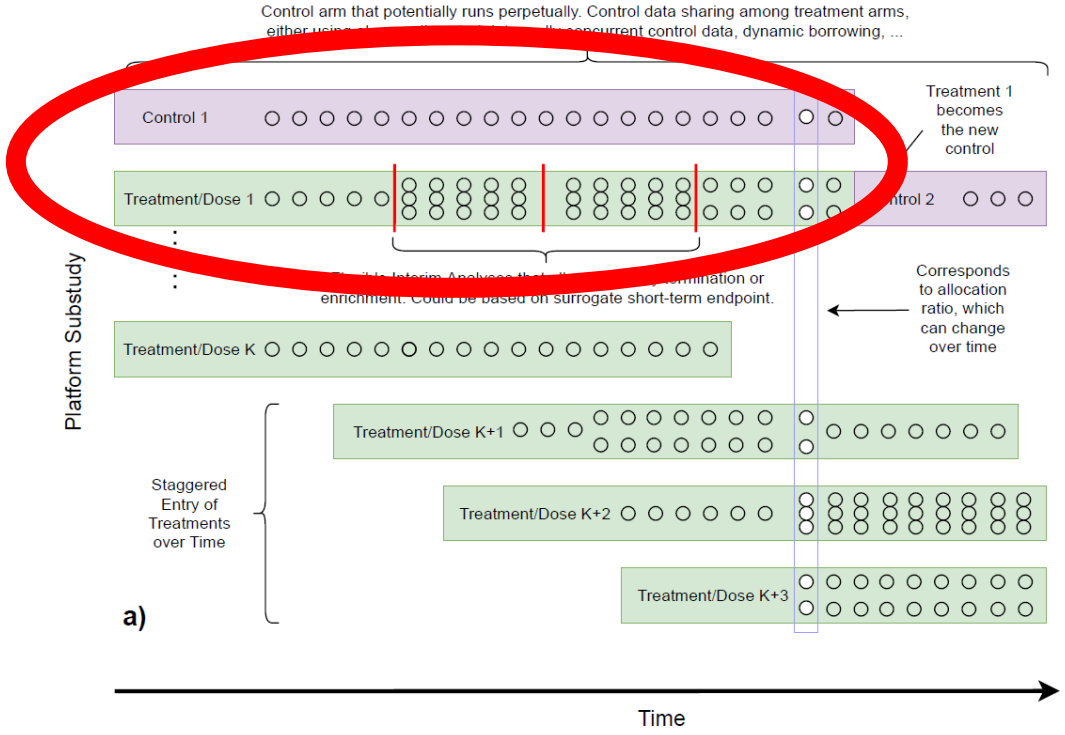
A pragmatic strategy for statistical inference

Each treatment/substudy in the platform trial is considered as an independent separate substudy, each controlling the FWER for the family of hypotheses relating to the treatment/substudy

For each substudy adjust for

- Multiple endpoints
- Multiple doses/treatment regimens
- Multiple subgroups
- Interim Analyses

But no adjustment across substudies



Summary Multiplicity in Platform Trials

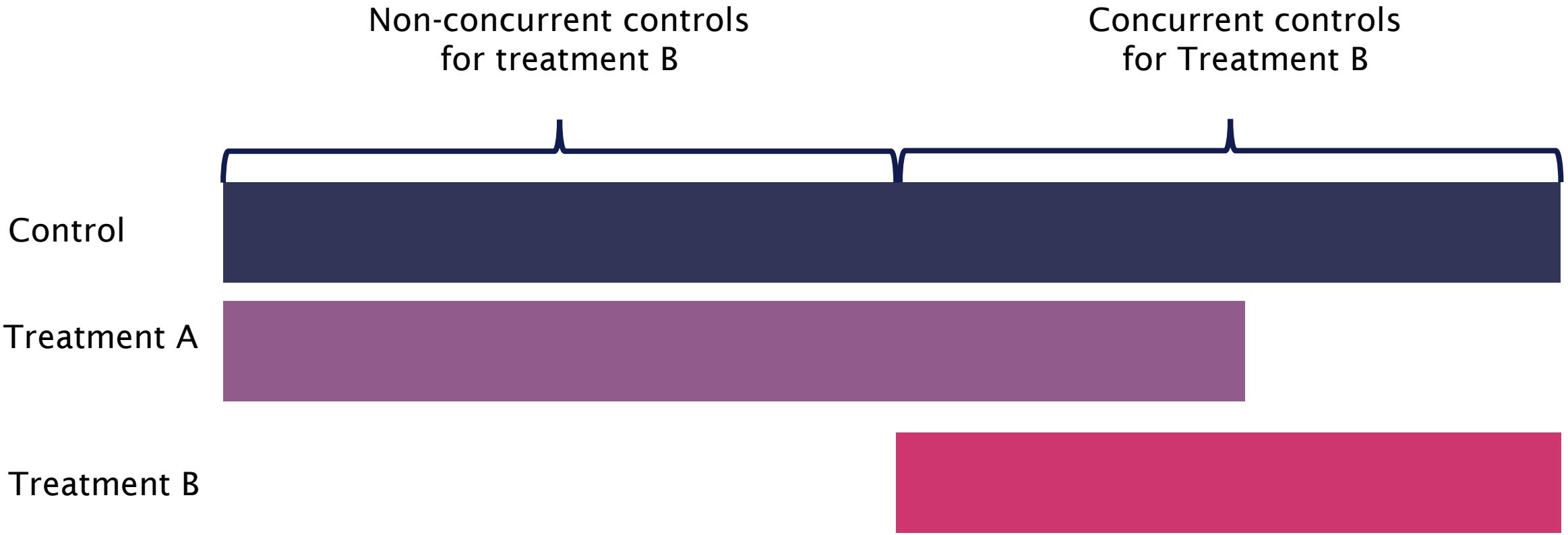
- The concept of study-wise T1E rate control is not directly applicable to platform trials, especially if they are perpetual in nature.
- Control of the Familywise Error Rate (FWER) rate at treatment or substudy level seems to be a pragmatic approach.
- But is there a consensus on what to consider “independent“?

- Also the overall operating characteristics of the platform trial are of importance. Depending on the trial objective, control of the FDR or FWER (possibly at higher levels) are possible options.
  - Online FDR Control: Zehetmayer, S., Posch, M., & Koenig, F. (2022). Online control of the False Discovery Rate in group-sequential platform trials. *Statistical Methods in Medical Research*, *31*(12), 2470-2485. Robertson, D. S., Wason, J. M., König, F., Posch, M., & Jaki, T. (2023). Online error rate control for platform trials. *Statistics in medicine*, *42*(14), 2475-2495.

- Other sources of multiplicity (treatments, change of control arms, subgroups, multiple endpoints, interim analysis, adaptations…) and sources of bias (non-concurrent controls, adaptations) need to be taken into account.

Shared and Non-Concurrent Controls

Can we use ALL control data, which is ALREADY available?



- If platform trials run over a long time period, with multiple treatments entering and leaving the platform over time, incorporating non-concurrent controls can substantially improve the efficiency
- However, non-concurrent controls may introduce bias due to different types of time trends

Non-Concurrent controls = Historical controls in RCT?

Non-concurrent and historical controls share several sources of potential bias
When using historical data for comparisons in clinical trials we accept that strict T1E control is not possible.

Eichler et al. 2016

So in platform trials?

Non-concurrent controls…

- are collected within a framework which has many features standardized (same infrastructure, assessment of endpoints, monitoring, …) and all changes are well documented.
- patients are randomized and blinding is possible

Randomized controlled trials & non-concurrent controls

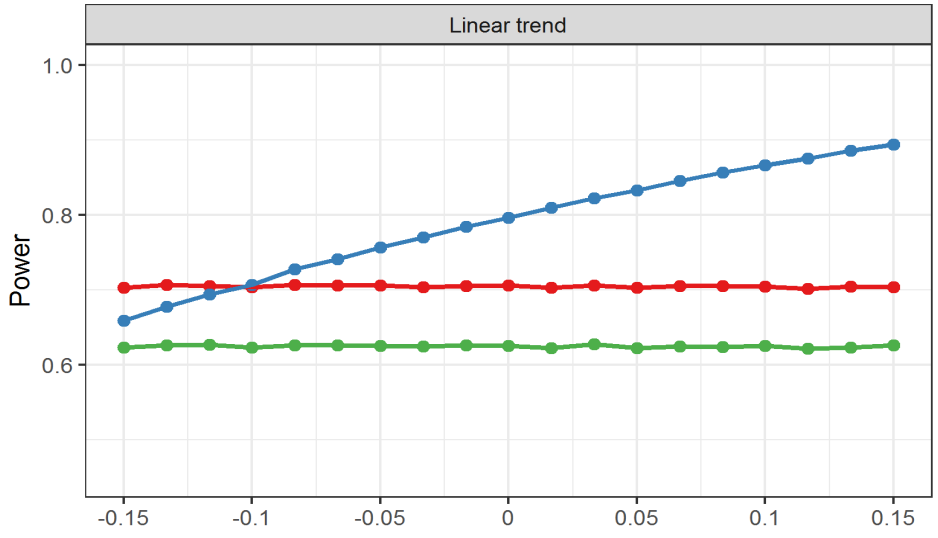
- Non-concurrent controls can be randomized & blinded but
  - At a **different calendar tim**e such that randomization **does not ensure control on the distribution of prognostic factors** between NCC and experimental arms.
  - patients & investigators **are not blinded with respect to the experimental treatment and the non-concurrent control** it is compared to
- The lack of true randomization can induce time trends

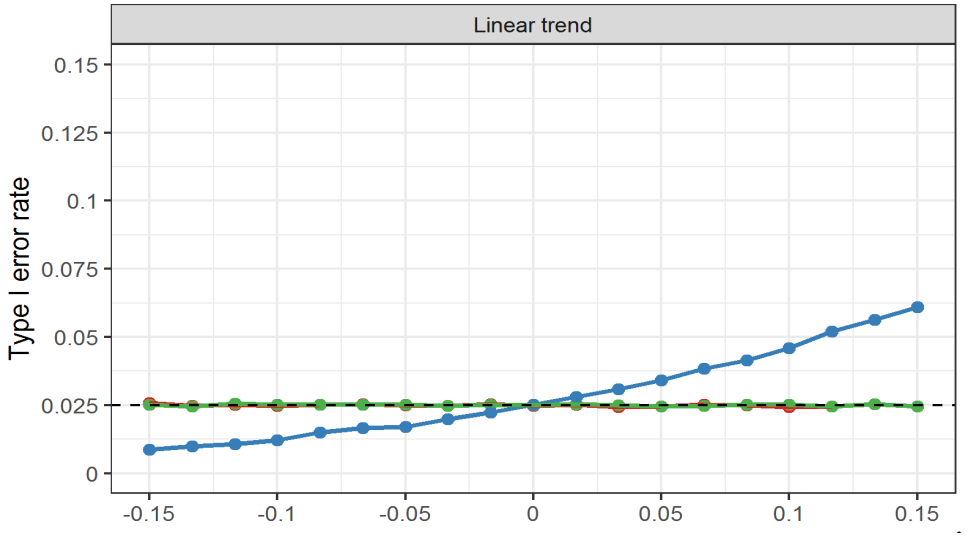
Time Trends due to External and Internal Factors

- **External,** e.g.,
  - Changes in standard of care
  - Patient population
  - Pandemics
- **Internal**
  - Change in **recruiting centers**: an analysis stratified by center is no longer possible if centers enter or leave the platform.
  - Change in **recruitment strategies**, e.g. if promising treatments enter the platform.
  - Change in **inclusion/exclusion criteria** because of other experimental treatments under investigation
  - Change in **assessment of endpoints** (e.g., new diagnostic devices)

Can we use all data?
Problem: Naively pooling control data can lead to error!

Example: 2 experimental arms and a control
Power and type 1 error rate as function of the strength of the linear time trend





- **Separate analysis** using only concurrent controls
- **Pooled analysis** using concurrent and non-concurrent controls
- **Regression model** adjusts for time trends in the model

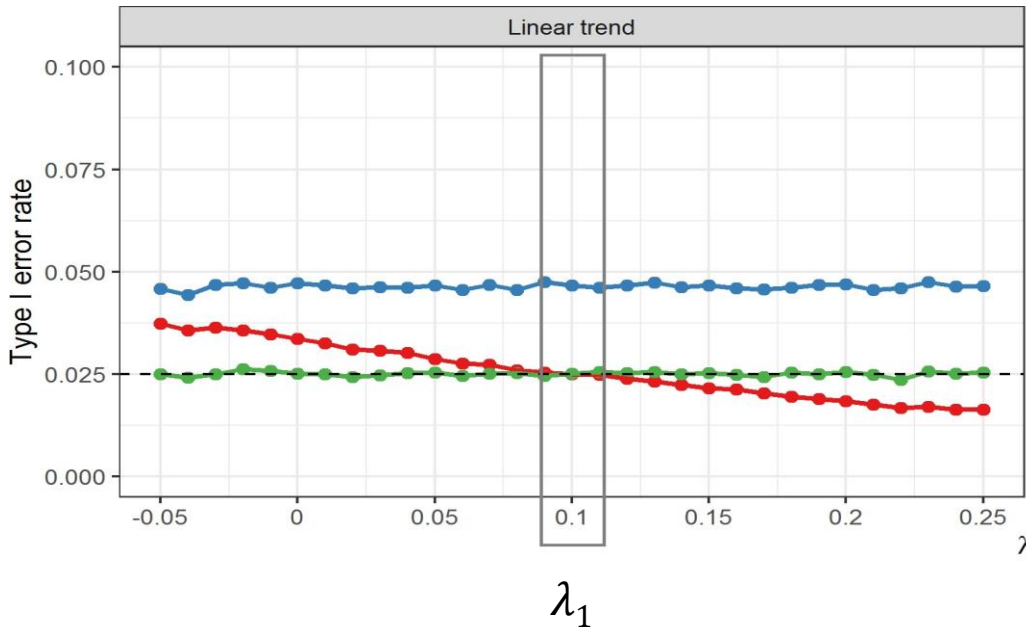
A solution: Bofill et al. (2022):
unbiased treatment effects regardless of the functional form of the time trend if
time trends in all treatment arms are equal and time trends are additive

EU-PEARL webinar:
https://eu-pearl.eu/workshops/non-concurrent-controls-in-platform-trials/
https://www.youtube.com/watch?v=nYl-lHtVwxA

Analysis approach: Regression model Pooled analysis Separate analysis

T1E for treatment arm 2 (different time trends in groups 1 and 2)

T1E as function of the strength of the time trend λ_1 in arm 1:



- **Separate analysis** using only concurrent controls
- **Pooled analysis** using concurrent and non-concurrent controls
- **Regression model** adjusts for time trends in the model

However, if time trends differ between treatment arms, estimates may be biased and the type 1 error rate may be inflated.

Bofill Roig, M. B., Krotka, P., Burman, C. F., Glimm, E., Gold, S. M., Hees, K., ... & Posch, M. (2022). On model-based time trend adjustments in platform trials with non-concurrent controls. BMC Medical Research Methodology, 22(1), 1-16.

Summary non-concurrent controls

- Inclusion of non-concurrent controls is a question of variance – bias tradeoff.
- Methods to address potential bias are available, however, they rely on specific assumptions.
- The problem of (the lack of) pre-specification is difficult to address. Keeping control data blinded may not be possible if treatment arms are stopped and results are reported.
- In broader indications regulators might be reluctant to accept analysis using NCC as well, but in rare diseases more efficient to use NCC data
- If non-concurrent data are utilized as primary analysis, also the analysis using only concurrent control data should be presented (possibly with a relaxed significance level)

Examples Platform Trials

A platform trial for neurofibromatosis (EU-PEARL)

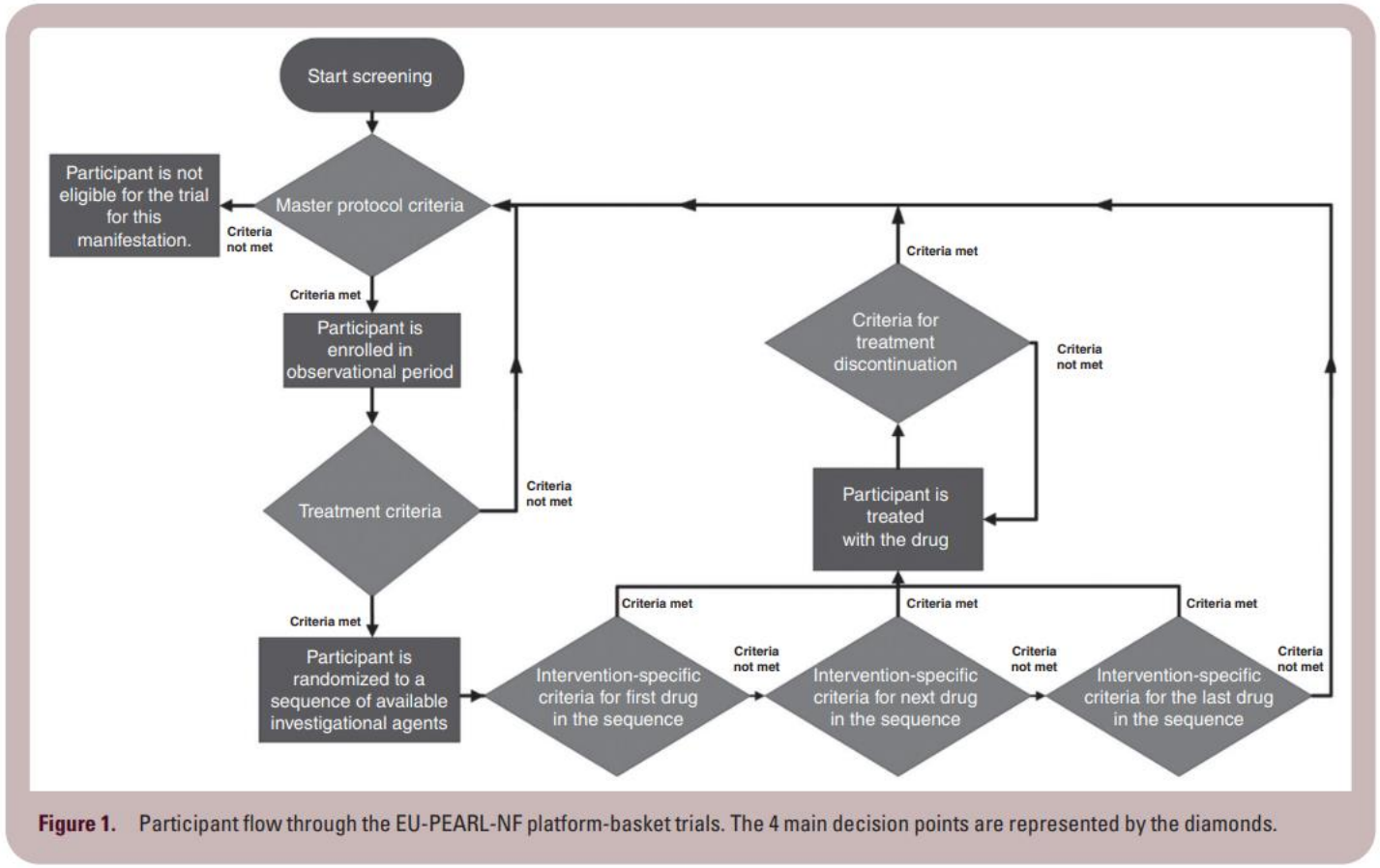


Figure 1. Participant flow through the EU-PEARL-NF platform-basket trials. The 4 main decision points are represented by the diamonds.

Neuro-Oncology Practice

1

XX(XX), 1–9, 2024 | https://doi.org/10.1093/nop/npae001 | Advance Access date 4 January 2024

Platform trial design for neurofibromatosis type 1, NF2-related schwannomatosis and non-NF2-related schwannomatosis: A potential model for rare diseases

Britt A.E. Dhaenens, Günter Heimann, Annette Bakker, Marco Nievo, Rosalie E. Ferner, D. Gareth Evans, Pierre Wolkenstein, Jonas Leubner, Cornelia Potratz, Charlotte Carton, Uchenna Iloeje, George Kirk, Jaishri O. Blakeley, Scott Plotkin, Michael J. Fisher, AeRang Kim, Pablo Hernáiz Driever, Amedeo A. Azizi, Brigitte C. Widemann, Andrea Gross, Tom Parke, Eric Legius, and Rianne Oostenbrink

All author affiliations are listed at the end of the article

Corresponding Author: Britt A.E. Dhaenens, MD, Department of General Paediatrics, Erasmus MC-Sophia Children's Hospital, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands (b.dhaenens@erasmusmc.nl).

Abstract

Background. Neurofibromatosis type 1, *NF2*-related schwannomatosis and non-*NF2*-related schwannomatosis (grouped under the abbreviation "NF") are rare hereditary tumor predisposition syndromes. Due to the low prevalence, variability in the range, and severity of manifestations, as well as limited treatment options, these conditions require innovative trial designs to accelerate the development of new treatments.

Methods. Within European Patient-Centric Clinical Trial Platforms (EU-PEARL), we designed 2 platform-basket trials in NF. The trials were designed by a team of multidisciplinary NF experts and trial methodology experts.

Results. The trial will consist of an observational and a treatment period. The observational period will serve as a longitudinal natural history study. The platform trial design and randomization to a sequence of available interventions allow for the addition of interventions during the trial. If a drug does not meet the predetermined efficacy endpoint or reveals unacceptable toxicities, participants may stop treatment on that arm and re-enter the observational period, where they can be re-randomized to a different treatment arm if eligible. Intervention-specific eligibility criteria and endpoints are listed in intervention-specific-appendices, allowing the flexibility and adaptability needed for highly variable and rare conditions like NF.

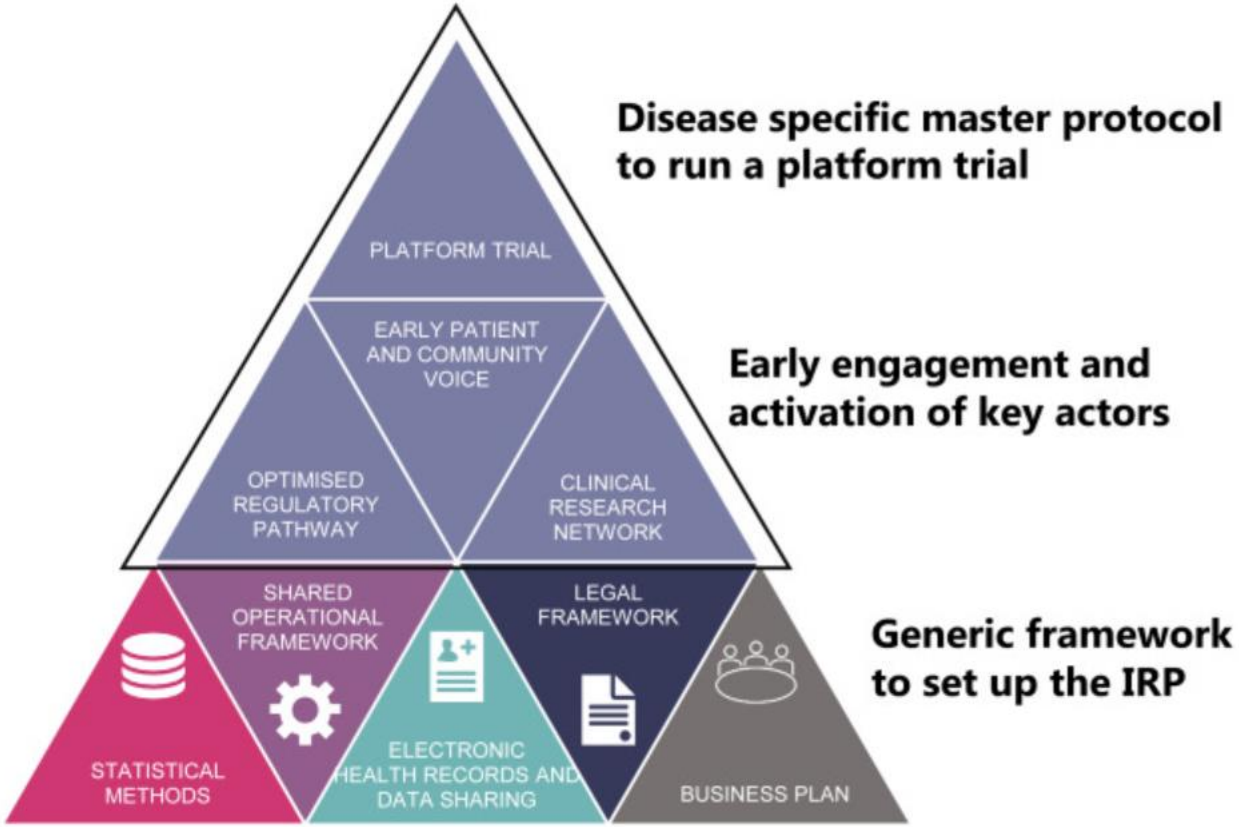
Conclusions. These innovative platform-basket trials for NF may serve as a model for other rare diseases, as they will enhance the chance of identifying beneficial treatments through optimal learning from a small number of patients. The goal of these trials is to identify beneficial treatments for NF more rapidly and at a lower cost than traditional, single-agent clinical trials.

Keywords

neurofibromatosis | schwannomatosis | clinical trial | platform trial | rare diseases

https://doi.org/10.1093/nop/npae001

Learnings on platform trials



ELEMENTS OF AN INTEGRATED RESEARCH PLATFORM

Current state-of-the-art and gaps in platform trials: 10 things you should know, insights from EU-PEARL

Franz Koenig,[a,j] Cécile Spiertz,[b,j] Daniel Millar,[c,j] Sarai Rodríguez-Navarro,[d] Núria Machín,[e] Ann Van Dessel,[f] Joan Genescà,[d,g,h,i] Juan M. Pericàs,[d,g,h,i,*,k] and Martin Posch,[a,k] on behalf of the EU-PEARL Consortium[l]

[a]Medical University of Vienna, Center for Medical Data Science, Vienna, Austria
[b]Janssen Biologics BV, Leiden, the Netherlands
[c]Former Employee, Janssen Research & Development, LLC, Raritan, NJ, USA
[d]Vall d'Hebron Institute for Research, Barcelona, Spain
[e]TEAM-IT Research S.L., Barcelona, Spain
[f]Janssen Pharmaceutica NV, Beerse, Belgium
[g]Liver Unit, Vall d'Hebron University Hospital, Barcelona, Spain
[h]Spanish Network of Biomedical Research Centers, Digestive and Liver Diseases (CIBERehd), Madrid, Spain
[i]Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Summary

Platform trials bring the promise of making clinical research more efficient and more patient centric. While their use has become more widespread, including their prominent role during the COVID-19 pandemic response, broader adoption of platform trials has been limited by the lack of experience and tools to navigate the critical upfront planning required to launch such collaborative studies. The European Union-Patient-cEntric clinicAl tRial pLatform (EU-PEARL) initiative has produced new methodologies to expand the use of platform trials with an overarching infrastructure and services embedded into Integrated Research Platforms (IRPs), in collaboration with patient representatives and through consultation with U.S. Food and Drug Administration and European Medicines Agency stakeholders. In this narrative review, we discuss the outlook for platform trials in Europe, including challenges related to infrastructure, design, adaptations, data sharing and regulation. Documents derived from the EU-PEARL project, alongside a literature search including PubMed and relevant grey literature (e.g., guidance from regulatory agencies and health technology agencies) were used as sources for a multi-stage collaborative process through which the 10 more important points based on lessons drawn from the EU-PEARL project were developed and summarised as guidance for the setup of platform trials. We conclude that early involvement of critical stakeholder such as regulatory agencies or patients are critical steps in the implementation and later acceptance of platform trials. Addressing these gaps will be critical for attaining the full potential of platform trials for patients.

eClinicalMedicine 2024;67: 102384
Published Online 26 December 2023
https://doi.org/10.1016/j.eclinm.2023.102384

Funding **Innovative Medicines Initiative 2 Joint Undertaking with support from the European Union's Horizon 2020 research and innovation programme and EFPIA.**

Keywords: **Adaptive designs; Master protocols; Patient-centred; Clinical research; Integrated research platform**

Introduction

Platform trials are increasingly used in clinical research and drug development in particular.[1] They are a form of adaptive design clinical trials that allow testing multiple interventions simultaneously and adding new treatments as they become available in the same trial structure within multiple subtrials (developed through intervention specific appendices-ISA), which can be either added or discontinued based on the results of interim analyses. Platform trials benefit from sharing trial infrastructure and resources, e.g., by sharing control data and joint committees. These elements are crucial in boosting both (1) the efficiency (i.e., the chances that a particular compound is graduated to the next clinical drug development phase if kept after an interim analysis, which also reduces the costs associated to carry on with the investment on a non-promising compound) and (2) the benefits for participants—i.e., increases the chances of participants being allocated to an intervention rather than placebo and the likelihood of receiving efficacious drugs as the platform trial progresses through interim analyses (see Fig. 1).

*Corresponding author. Liver Unit, Vall d'Hebron University Hospital, Main Building, Ground Floor, Desk 071, 08035, Barcelona, Spain.
E-mail address: juanmanuel.pericas@vallhebron.cat (J.M. Pericàs).
[j]Contributed equally.
[k]Co-senior authors.
[l]The complete list of EU-PEARL' investigators can be found in the Supplementary Material.

DOI:https://doi.org/10.1016/j.eclinm.2023.102384

Further papers on specific designs and methodology for platform trials, e.g., multiplicity, use of non-concurrent controls

Templates to create a Master Protocol

- The **Master protocol** governs the entire study and includes the **common key study design elements**
- **Intervention-specific information** is provided in **Intervention Specific Appendices (ISAs)**, which are added as interventions become available and are ready to enter the platform study
- Interventions can enter the platform study **simultaneously** or **sequently** as they become available for study
- **Both protocols** are needed to have all the information needed to conduct the study in an intervention cohort

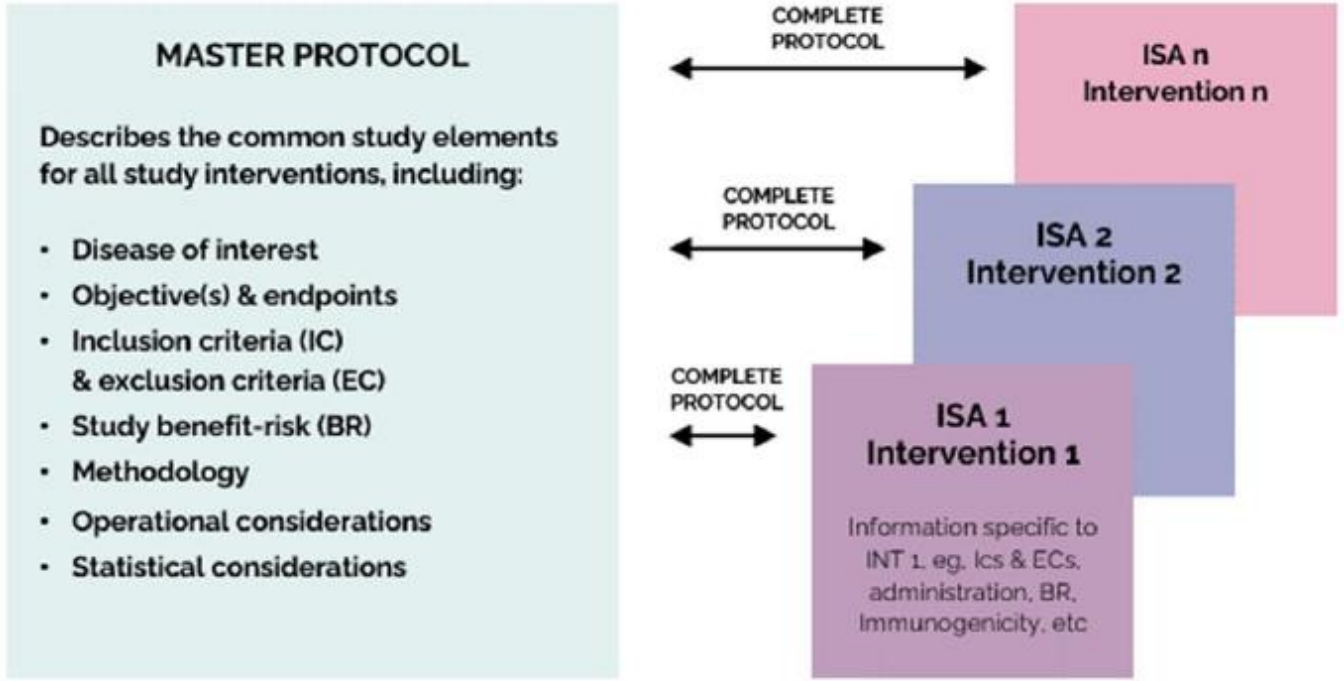


Fig. 2 Outline of the master protocol and ISA or sub-protocol

I. Provisional Master Protocol Template
II. Provisional Intervention Specific Appendix (ISA)
III. Provisional Statistical Analysis Plan Template
IV. Provisional Data Monitoring Charter Template
Platform Trial Best Practises Tool

Reference: Mesenbrink, P., Gidh-Jain, M., Parke, T., Koenig, F., & Spiertz, C. (2023). Developing Generic Templates to Shape the Future for Conducting Integrated Research Platform Trials. Pre-print at https://doi.org/10.21203/rs.3.rs-3382348/v1 Trials 2024

Software for adaptive designs and master protocols

Meyer et al. Trials (2021) 22:183
https://doi.org/10.1186/s13063-021-05130-x

Trials

REVIEW Open Access

Systematic review of available software for multi-arm multi-stage and platform clinical trial design

Elias Laurin Meyer[1], Peter Mesenbrink[2], Tobias Mielke[3], Tom Parke[4], Daniel Evans[5], and Franz König[1*] on behalf of EU-PEARL (EU Patient-cEntric clinicAl tRial pLatforms) Consortium

Abstract

Background: In recent years, the popularity of multi-arm multi-stage, seamless adaptive, and platform trials has increased. However, many design-related questions and questions regarding which operating characteristics should be evaluated to determine the potential performance of a specific trial design remain and are often further complicated by the complexity of such trial designs.

Methods: A systematic search was conducted to review existing software for the design of platform trials, whereby multi-arm multi-stage trials were also included. The results of this search are reported both on the literature level and the software level, highlighting the software judged to be particularly useful.

Results: In recent years, many highly specialized software packages targeting single design elements on platform studies have been released. Only a few of the developed software packages provide extensive design flexibility, at the cost of limited access due to being commercial or not being usable as out-of-the-box solutions.

Conclusions: We believe that both an open-source modular software similar to OCTOPUS and a collaborative effort will be necessary to create software that takes advantage of and investigates the impact of all the flexibility that platform trials potentially provide.

Introduction

Master protocol trials allow for the evaluation of both multiple investigational treatments and multiple subgroups of the study population within the same overall clinical trial structure, as compared to traditional randomized controlled trials, where usually only one investigational treatment is investigated in one study population [1]. Several types of master protocol trials can be distinguished, such as basket trials, umbrella trials, and platform trials. Whereas in classical development programs different studies are needed for newly available treatments, adaptive platform trials are a type of randomized clinical study that allow for the evaluation of multiple interventions in a disease or condition in a perpetual manner, with interventions entering and leaving the platform on the basis of a predefined decision algorithm (definition following the Adaptive Platform Trials Coalition [2]). Figure 1 illustrates the difference between the platform paradigm and a classical drug development program. One of the major advantages of platform trials is their reduced sample size due to the sharing of a common control arm. The platform trial design offers other important potential advantages compared to the traditional approach of running many studies either sequentially or in parallel, including an overall reduction in the trial infrastructure and the removal of competition between trials within a limited pool of

* Correspondence: franz.koenig@meduniwien.ac.at
[1]Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria
Full list of author information is available at the end of the article

https://doi.org/10.1186/s13063-021-05130-x

SoftwareX 23 (2023) 101515

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

SIMPLE—A modular tool for simulating complex platform trials

Elias Laurin Meyer[a,b], Tobias Mielke[c], Tom Parke[d], Peter Jacko[d,e], Franz Koenig[a,*]

[a] Center for Medical Data Science, Medical University of Vienna, Austria
[b] Berry Consultants, Vienna, Austria
[c] Janssen Cilag GmbH, Germany
[d] Berry Consultants, Abingdon, UK
[e] Lancaster University, Lancaster, UK

ARTICLE INFO

Article history:
Received 4 April 2023
Received in revised form 11 July 2023
Accepted 24 August 2023

Keywords:
R
Data visualization
Platform trials
Clinical trial simulation
Adaptive trials

ABSTRACT

Platform trials are becoming increasingly popular within drug development, attracting interest by patients, clinicians, regulatory agencies and statisticians. More often than not, these platform trial designs are highly complex and involve many weakly predictable events (e.g. number of investigational treatments that will enter over time) to determine the impact of relevant design parameters (e.g. decision rules, sharing of information across cohorts and allocation ratios) on the operating characteristics with high confidence. Simulations may address these uncertainties at the design stage. However, the number and combination of design elements for potential implementation in real platform trials is immense. As a result, simulation software which is developed based on specific project needs is typically limited in the variety of available design options for comparison, as such software is developed for a particular need, not for researching all potential new approaches to clinical research and statistical science. On the other hand, software solutions which allow for a wide range of design options may easily overload the user with requirements for design specifications. We have developed an R software package ("**SIMPLE**"), which is modular in the sense that if users want to simulate the most common platform designs, minimal effort and understanding of the package is needed, but it allows the users to take control of different parts of the simulation (e.g. patient accrual, outcome simulation, etc.) step-by-step, thereby facilitating the simulation of arbitrarily complex platform trials. We will give an overview of this software package alongside some examples on how to simulate common platform trial designs and derive their operating characteristics.

https://doi.org/10.1016/j.softx.2023.101515

EU-PEARL: Simulation software in **public domain** (EU-PEARL webpage, publications, open source software simulation programs https://github.com/EUPEARL, ...)

RESEARCH ARTICLE

Designing an exploratory phase 2b platform trial in NASH with correlated, co-primary binary endpoints

Elias Laurin Meyer[1], Peter Mesenbrink[2], Nicholas A. Di Prospero[3], Juan M. Pericàs[4,5], Ekkehard Glimm[6,7], Vlad Ratziu[8], Elena Sena[4], Franz König[1]*, on behalf of the EU-PEARL NASH Investigators[¶]

Original software publication

NCC: An R-package for analysis and simulation of platform trials with non-concurrent controls

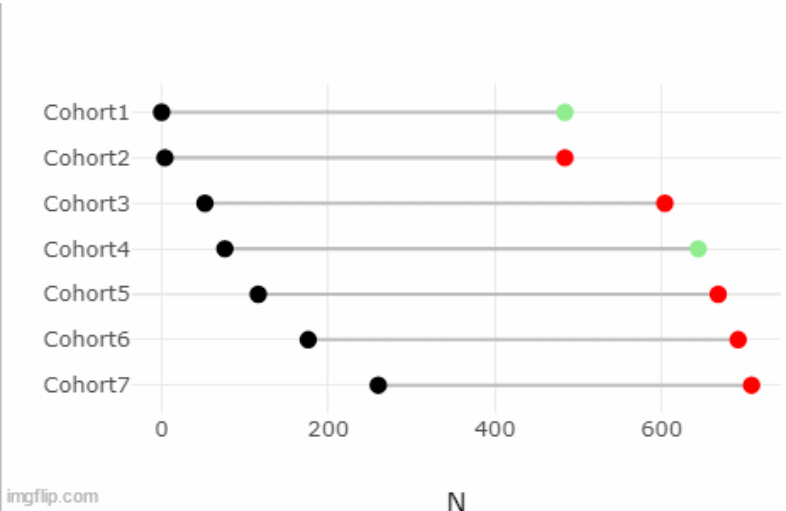
Pavla Krotka[a], Katharina Hees[b], Peter Jacko[c,d], Dominic Magirr[e], Martin Posch[a], Marta Bofill Roig[a,*]

[a] Center for Medical Data Science, Medical University of Vienna, Vienna, Austria
[b] Section of Biostatistics, Paul-Ehrlich-Institut, Langen, Germany
[c] Berry Consultants, Abingdon, UK
[d] Lancaster University, Lancaster, UK
[e] Advanced Methodology and Data Science, Novartis Pharma AG, Basel, Switzerland

Conclusions

Conclusions (I) Adaptive Clinical Trials

- Better use of resources versus traditional parallel group design as they allow for mid-trial learning and adaptations while strictly controlling the (multiple) type I error rate.
- Platform trials are a challenge and opportunity
- Analytic solutions to evaluate OCs (T1E, power) often not available
- Increased use of clinical trial simulations



Meyer, E. L., Mielke, T., Parke, T., Jacko, P., & König, F. (2023). SIMPLE—A modular tool for simulating complex platform trials. SoftwareX, 23, 101515.
Meyer, Bofill-Roig, Jacko, Krotka, Mesenbrink, Zehtmayer, Zocholl, König. Why and how should we simulate platform trials? - Learnings from EU-PEARL (2024). Submitted
Krotka, P., Hees, K., Jacko, P., Magirr, D., Posch, M., & Roig, M. B. (2023). NCC: An R-package for analysis and simulation of platform trials with non-concurrent controls. SoftwareX, 23, 101437.
Meyer, E. L., Mesenbrink, P., Mielke, T., Parke, T., Evans, D., König, F., (2021). Systematic review of available software for multi-arm multi-stage and platform clinical trial design. Trials, 22, 1-14.

Conclusions (II)

Questions that should generally be addressed during planning and assessment

1. Is there a good rationale? Have alternative, more standard trial designs been considered?
2. Does the proposal fit well in the context of the development program and the data that will be available for the marketing authorization application?
3. Can the proposal be implemented without important damage to trial integrity?
4. Is the type I error rate controlled?
5. Has the potential bias of treatment effect estimates been evaluated?
6. Is the proposal practical and feasible?

Conlusions (III)

- Adaptive designs seem well accepted if properly planned and implemented
- A range of increasingly complex adaptive designs are proposed, the majority in rare diseases
- Surprisingly, still a lack of methodological knowledge
  - how to achieve type I error control
  - how to assess the efficiency of the design (timing of interim analysis, adaptation rules, power)
- Who should be decide on adaptations at interim, (DMC?, sponsor?, ...)
- Group sequential designs developed in the 70s are now well established - do we still have to wait one decade until the adaptive methodology is common knowledge?

Selected References (see references in the three papers)

- Bauer, P., Bretz, F., Dragalin, V., König, F., & Wassmer, G. (2016). Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls. Statistics in Medicine, 35(3), 325-347.
- Koenig, F., Spiertz, C., et al . (2024). Current state-of-the-art and gaps in platform trials: 10 things you should know, insights from EU-PEARL. The Lancet eClinnicalmedicine, 67. https://doi.org/10.1016/j.eclinm.2023.102384
- Meyer, E. L., Mesenbrink, P., ... & König, F. (2020). The evolution of master protocol clinical trial designs: a systematic literature review. Clinical Therapeutics, 42(7), 1330-1360.

Acknowledgements (for part on platform trials)

Thanks to Martin Posch, Elias Meyer, Marta Bofill Roig, Sonja Zehetmayer, Peter Mesenbrink, Ceciele Spietz, Ekkehard Glimm, Tobias Mielke, Tom Parke, Benjamin Hofner, Pavla Krotka, and all other EU-PEARL members

https://eu-pearl.eu/

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 853966.
The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and CHILDREN'S TUMOR FOUNDATION, GLOBAL ALLIANCE FOR TB DRUG DEVELOPMENT NON PROFIT ORGANISATION, SPRINGWORKS THERAPEUTICS INC.

Many thanks for your patience