Basel Biometric Society Course on Data Monitoring Committees (DMCs)

Meet the presenters



Tim Friede



David Lawrence



Tobias Mütze



Marisa Bacchi

Disclaimer

The opinions and views expressed in this slide deck are those of the presenters. Each presenter speaks in an individual capacity, and their contributions do not necessarily reflect the views or recommendations of their employer.

House keeping

- Please ask questions at any time during the course.
- Use the breaks to introduce yourselves.
- Share your experience preparing outputs for DMCs, working with DMCs, or serving on a DMC.
- Be ready to learn from one another and share your knowledge.

Agenda

DAY 1 (13:00-18:00)

Welcome

Basic principles of DMC

How DMCs are operating

Practical 1: Open session

Coffee break (30’)

DMC charter as a governing document

Practical 2: Communication between stakeholder

Q&A

DAY 2 (8:30-12:00)

DMC reports and reporting

Coffee break (30')

Overview of statistical methods for monitoring

DMC's in non-standard settings

Q&A

Basic principles for DMCs

Tim Friede

Many names of data monitoring committees

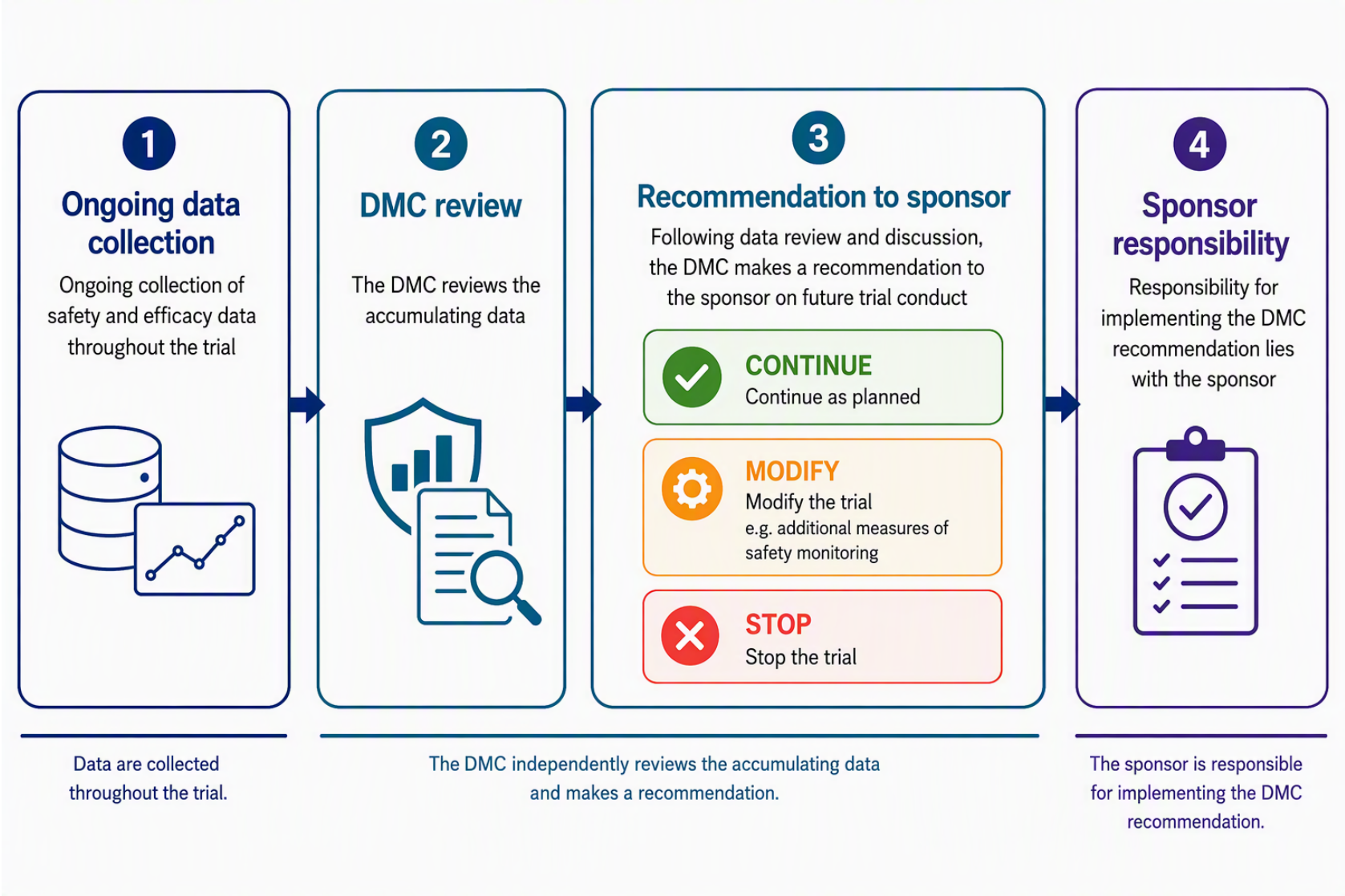
- Data monitoring committees (DMC) also known as
  - data safety monitoring board (DSMB)
  - data safety monitoring committee (DSMC)
  - Sometimes “independent” (IDMC) is added to the name
- Both FDA and EMA use the term DMC in their guidance documents

What is the purpose of a DMC?

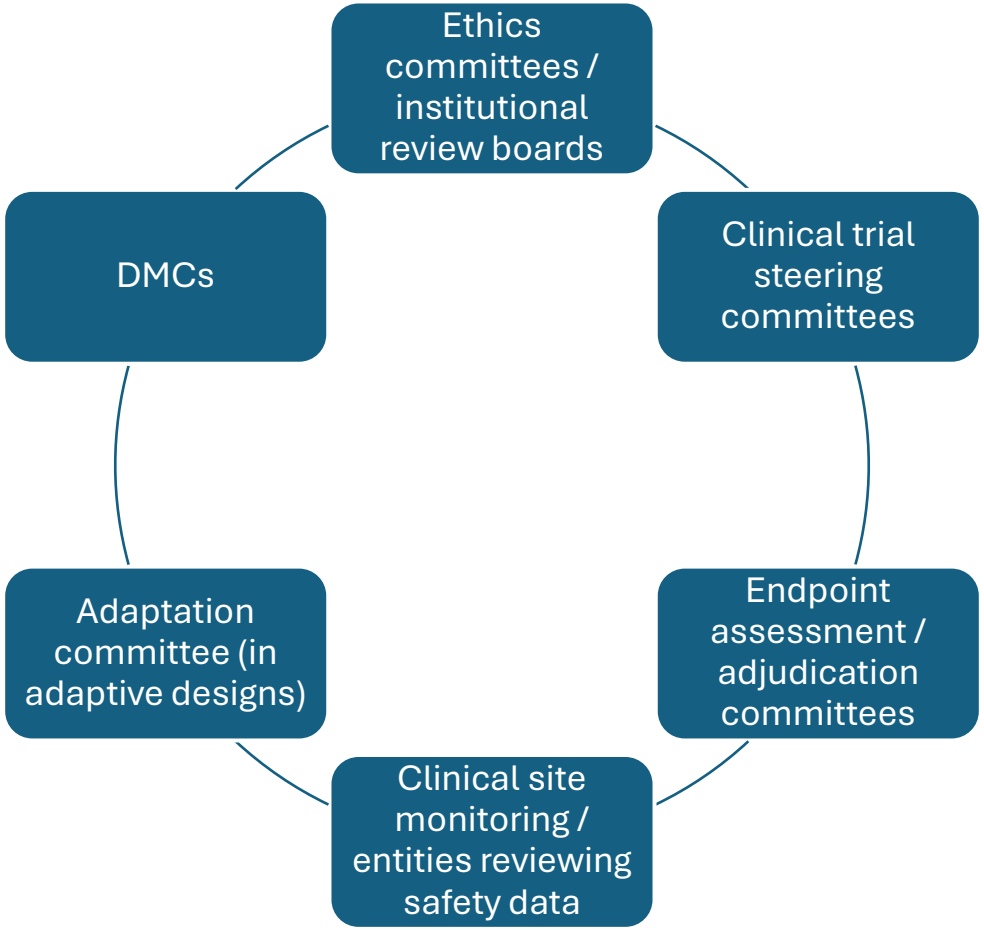
- “[…] a committee **created by a sponsor** to provide independent review of **accumulating safety and efficacy data**. This committee is responsible for the stewardship of the trial which translates to **protecting trial integrity and patient safety** […]”

Text from the Glossary in Herson, 2009; bold font used here for presentation

DMCs make recommendations to sponsor



DMCs is one out of a number of oversight groups monitoring trial progress including

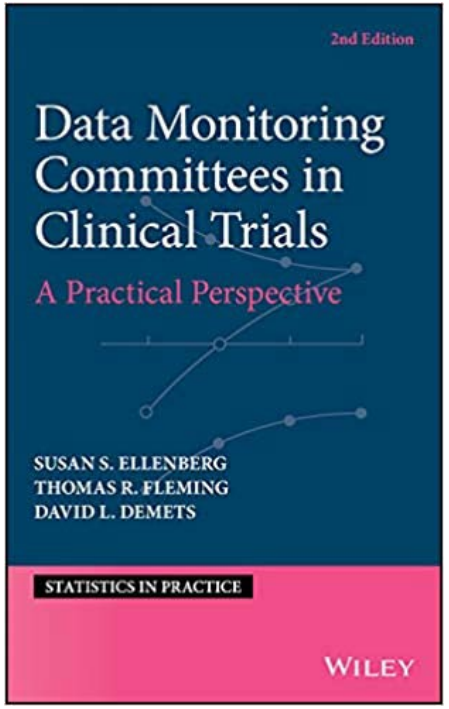


Some history . . .

- Practices involving DMCs evolved from the *Greenberg Report* (1967), charged by NIH to develop recommendations for clinical trial conduct
- Recommended that a board of experts not otherwise involved in the trial should review data as the trial progresses, and make recommendations as needed
- Practice evolved to exclude study leadership from interim data access
  - Maximally objective review
  - Avoided bias in conducting / managing the trial
- Over time, this model became more commonly implemented in trials supported by NIH, NCI, others

More recently…

- DMCs were uncommon in industry trials prior to mid-1990's
  - My sense: industry practices, both operationally and methodologically, were inconsistent through that time
- First FDA draft DMC guidance (Nov. 2001) was a big step forward
- Followed by the highly influential text by Ellenberg et al (2003)
  - {2nd edition: 2019)



When is a DMC needed? (FDA 2006 guidance)

- Consider implementing a DMC when …
  - “a highly favorable or unfavorable result [..] at an interim analysis might ethically require termination of the study”
  - prior information regarding safety concerns / toxicities exists
  - “fragile population such as children, pregnant women or the very elderly, or other vulnerable populations”
  - “population at elevated risk of death or other serious outcomes, even when the study objective addresses a lesser endpoint”
  - “The study is large, of long duration, and multi-center.”

- DMC is not needed for short term trials for symptom relief

When is a DMC needed? (FDA 2024 draft guidance)

- **Determining whether to use a DMC (Section IV)**
  - "An important consideration in determining whether to use a DMC in a development program is whether DMC review is practical for the particular clinical trial."
  - "FDA strongly recommends establishing a DMC if trial subjects are at risk of serious morbidity or mortality (e.g., hospitalization, heart attack, stroke, death)."
  - "Sometimes the DMC is used to make recommendations on operational matters based on accumulating noncomparative safety and efficacy data [...]. These findings can also be addressed by other groups (e.g., clinical trial steering committees)."

Composition of DMCs

- Essential DMC members
  - Clinicians with expertise in the medical area under study
  - Statistician experienced in clinical trials and statistical monitoring plans
- Additional expertise (as needed)
  - Bioethicists and patient advocates
  - Specialists in pharmacology, toxicology, behavioral science, etc.
- Experience requirements
  - Prior experience in clinical trials and, ideally, prior DMC participation
- Members should be
  - Independent of sponsor (more on internal DMCs later!)
  - Represent the regions patients are recruited from
  - Limited conflicts of interest
- Members are appointed by the sponsor (possibly after consultation with the steering committee)

Format of DMC meetings

- DMCs meet regularly while the trial is ongoing
- DMC meeting a comprised of an open and closed session
  - Open session: review of trial progress (e.g., recruitment, training of trial sites, data cleaning) and blinded safety data
  - Closed session: review of unblinded safety and efficacy data (restricted to DMC members and statistical analysis center)
  - Executive session (optional; restricted to DMC members only)
- Open feedback session with sponsor representatives (used to be common with F2F meeting; discouraged due to risk of unblinding!)
- DMC makes recommendation to sponsor on future trial conduct after the meeting
- More on this later…

Regulatory guidance on DMCs: Key principles are largely consistent

2005: Operational Guidelines for the Establishment and Functioning of Data and Safety Monitoring Boards

2006: Guideline on Data Monitoring Committees

2006: Guidance for Clinical Trial Sponsors: Establishment and Operation of Clinical Trial Data Monitoring Committees

2013: Guideline on Data Monitoring Committees

2020: Guideline on Clinical Trial Data Monitoring Committees

2020: Questions and Answers on Data Monitoring Committee Issues

2024: Use of Data Monitoring Committees in Clinical Trials: Guidance for Industry (draft)

Other guidelines (e.g., on safety reporting) might also comment on DMCs

DMCs in ICH guidelines

Guideline	Context	Link
E6 (R2/R3) – Good Clinical Practice	Defines DMC role and responsibilities in trial oversight	E6(R2) · E6(R3)
E3 – Structure and Content of Clinical Study Reports	Requires reporting DMC details in CSR	E3 Guideline
E9 – Statistical Principles for Clinical Trials	Discusses role of DMCs	E9 Guideline
E17 – Multiregional Clinical Trials	Mentions monitoring and oversight in MRCTs	E17 Guideline
E19 – Selective Safety Data Collection	Safety oversight context, implies DMC role	E19 Guideline
E20 (Draft) – Adaptive Design for Clinical Trials	Highlights DMC role in adaptive designs and interim analyses	E20 Draft

Group discussion

- How are DMC members chosen in your organization?

Basic principles for DMC membership

- Independence
- Conflicts of interest
- Confidentiality
- Indemnification

Independence from the sponsor

- No significant conflicts of interest
  - Detailed discussion on next slide
- Independent engagement
  - Contracted as “independent scientists,” not regular consultants.
  - Preferably appointed by an external steering committee.
- Transparent financial disclosure
  - Annual disclosure of consulting roles and research funding.
  - Reviewed by DMC chair or sponsor.
- No dual roles
  - Not involved as trial investigators.
  - No authorship expectations on primary trial publications.
- Structured oversight to preserve impartiality
  - E.g., through use of open/closed sessions and a clear DMC Charter.

Conflicts of interest (I)

- Various types of conflicts of interest exist: Financial, intellectual, professional/emotional
- Financial conflicts
  - receiving excessive honoraria or stock options from the sponsor
  - owning stock in the sponsor or competing companies
  - depending on sponsor for future research funding
  - being the inventor or developer of the product under evaluation
- Intellectual conflicts
  - having strong opinions or emotional investment in the intervention
  - prior involvement in developing or promoting the product
  - serving as investigator in a competing trial of a similar product
- Professional and emotional conflicts
  - close relationships with sponsor staff or trial investigators
  - desire to protect careers or reputations of colleagues
  - advocacy group members with strong emotional ties to outcomes

Conflicts of interest (II)

- Conflicts of interest should be evaluated when choosing individuals to serve on a DMC
- “Aside from being compensated for their duties on the committee, DMC members should have no ongoing financial relationship with a trial’s commercial sponsor (or its direct competitors) […]”
  - Quote from FDA draft DMC guidance
- How to mitigate conflicts of interests?
  - Annual disclosure of financial interests and consulting roles
  - Use of independent contracting and review processes

Confidentiality

- Confidentiality of (unblinded) interim trial results ensures trial integrity
- Correspondence to and from the DMC should be regulated (flow chart) and, DMC members should not disclose interim information from the study to the sponsor.
- The DMC Charter contains guidance on exchange of information and how to provide recommendations to the Sponsor without compromising data integrity
- The DMC members should be trained/aware of these aspects of confidentiality

When can DMC members be internal?

- Internal DMCs or hybrid DMCs are DMCs where (part of) the DMC members are employees of the sponsor
- External DMCs are recommended for confirmatory trials, or when health authorities require independent oversight
  - Ensures credibility and unbiased monitoring
- Internal or hybrid DMCs might be acceptable for exploratory studies focused on internal decision-making

Indemnification

- Indemnification refers to the protection of DMC members from legal liability arising from their roles in clinical trials
- Many DMC members are concerned about potential legal action taken by patients or investors
- Indemnification ensures that DMC members can perform their duties without fear of litigation
  - Maintains the integrity and independence of the committee
- DeMets et al Hess et al Filippatos et al DeMets & Ellenberg call for indemnification of DMC members by the sponsor
  - This should include support to cover legal counsel for the DMC member independent of the sponsor's legal counsel to avoid legal conflicts of interest

Indemnification

- Typically, indemnification is agreed on in the contract between DMC member and sponsor. An indemnification clause might read like this:
  - "SPONSOR shall indemnify, defend, and hold harmless Consultant from any and all third-party losses, costs, expenses, liabilities, claims, actions, lawsuits and damages including reasonable legal fees and expenses that may arise or result from the existence, performance and termination of the Agreement."
- When the contract is between the CRO providing services to the DMC and the DMC member, then the situation is more complex. Then the sponsor might issue a separate letter of indemnification.

How DMCs are operating

Tobias Mütze

Four questions guide this section

Who decides what?

Who does what?

Who sees what?

Who documents what?

Deciding whether to include a DMC

- Who decides?
  - Sponsor, but health authorities may request one
- When should decision be made?
  - Time of protocol development
- Where to document?
  - Protocol needs to mention that a DMC will monitor trial
  - ICH M11 protocol template good indicator what to document
- Good practice: send DMC members a final draft of the protocol (and amendments) for review

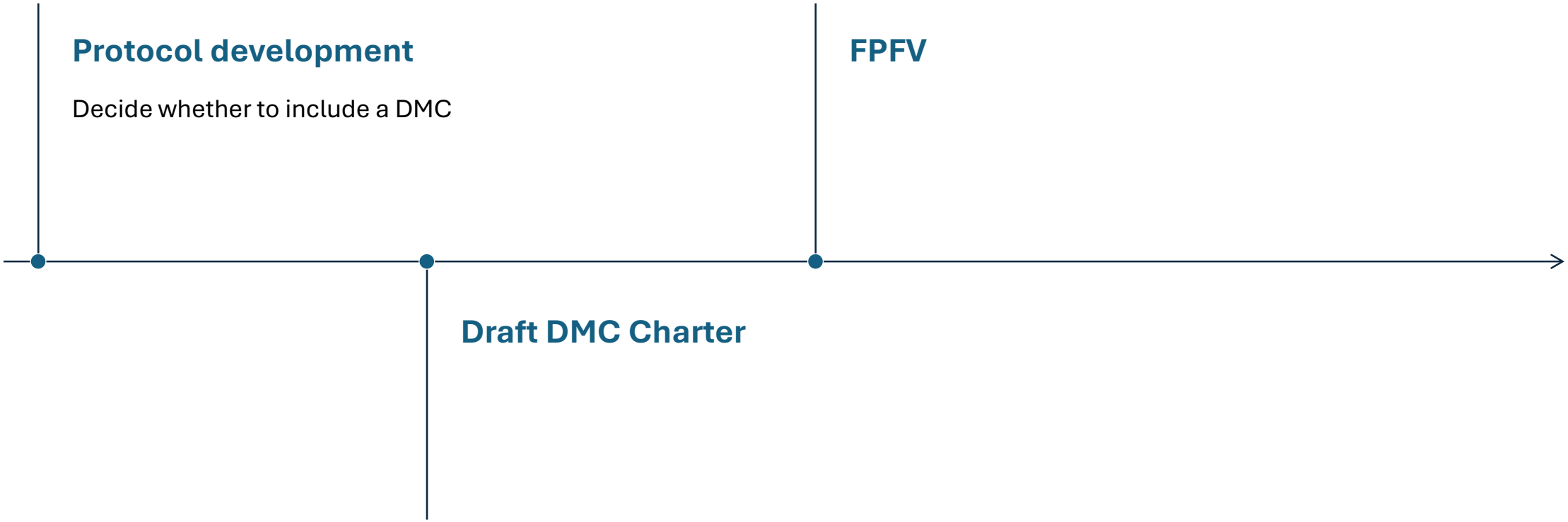
DMC timeline



Charter governs the DMC's work

- DMC charter is a formal written agreement that defines the roles, responsibilities, and operating procedures of the DMC throughout the clinical trial
- Draft FDA DMC guidance: "All DMCs should operate under a written charter that clearly states the purpose of the DMC, the specific questions it is expected to address, and the possible recommendations it can make to the sponsor during the trial."
- Completed and signed prior to first patient first visit
  - ICH E8(R1): "If a data monitoring committee is needed for either an individual study or across the development programme, procedures governing its operation [...] should be established prior to study start."
- More later...

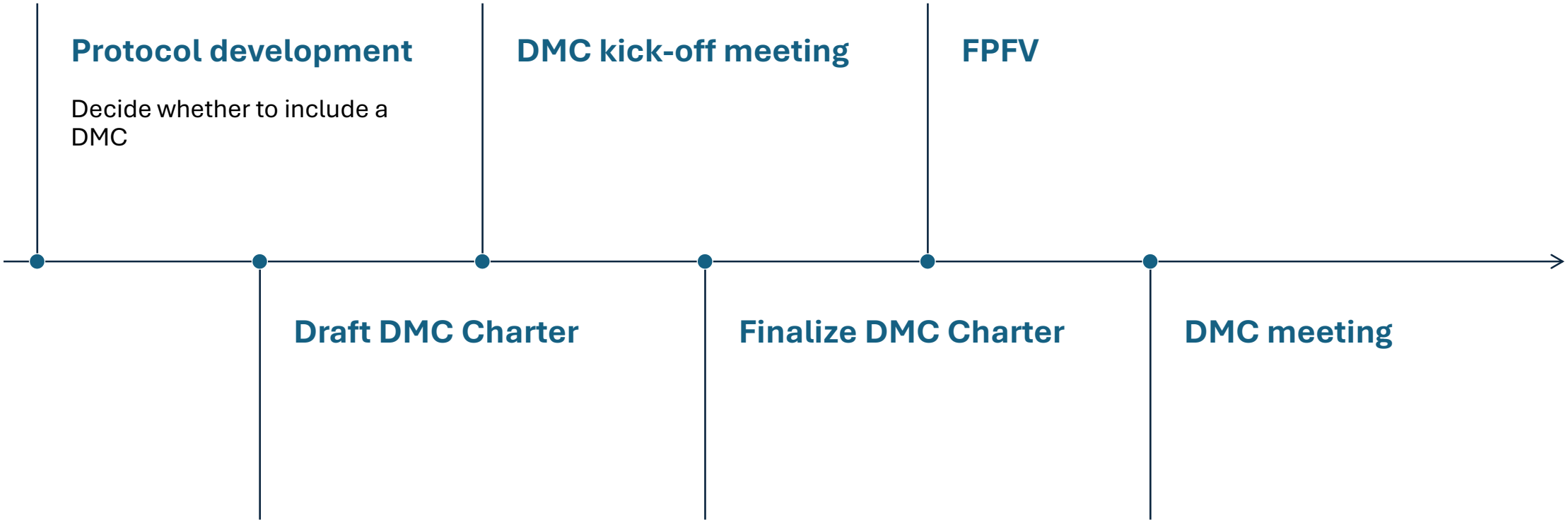
DMC timeline



DMC kick-off meeting

- Meeting to align on various organizational aspects and to give the DMC an understanding of the trial expectations from the Sponsor
- Topics at the kick-off meeting
  - Review trial design and conduct
  - Discussion of draft DMC charter
  - Contents of closed DMC report
  - Review of efficacy/futility stopping boundaries
  - Introduction of the independent statistical team
  - …
- Must be held prior to first patient, first visit

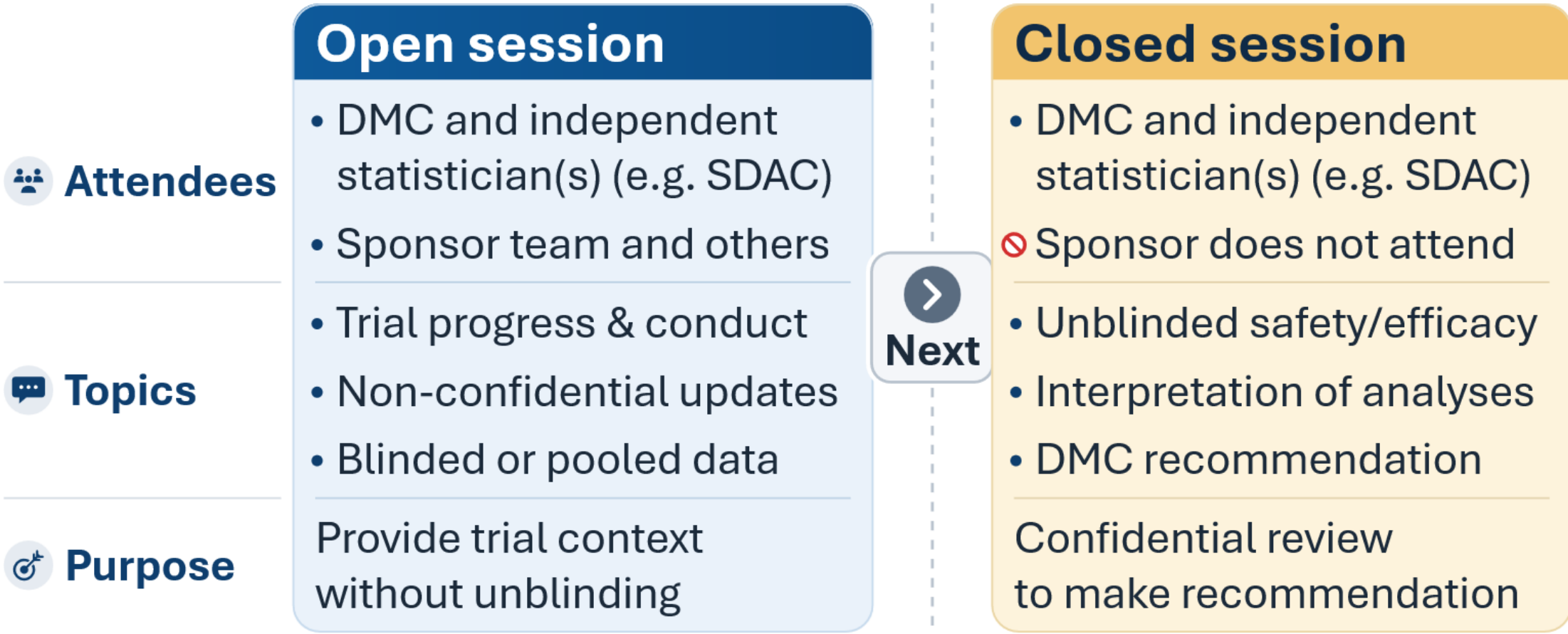
DMC timeline



DMC meetings and their preparation

- Overall goal: DMC to make a recommendation to the sponsor on future trial conduct based on review of unblinded data
- Frequency of meetings depends on the trial, but default is twice a year
- DMC meetings are split into an open and a closed session
- Who prepares analysis of unblinded data?
  - SDAC: Statistical data analysis center
  - More on this later…

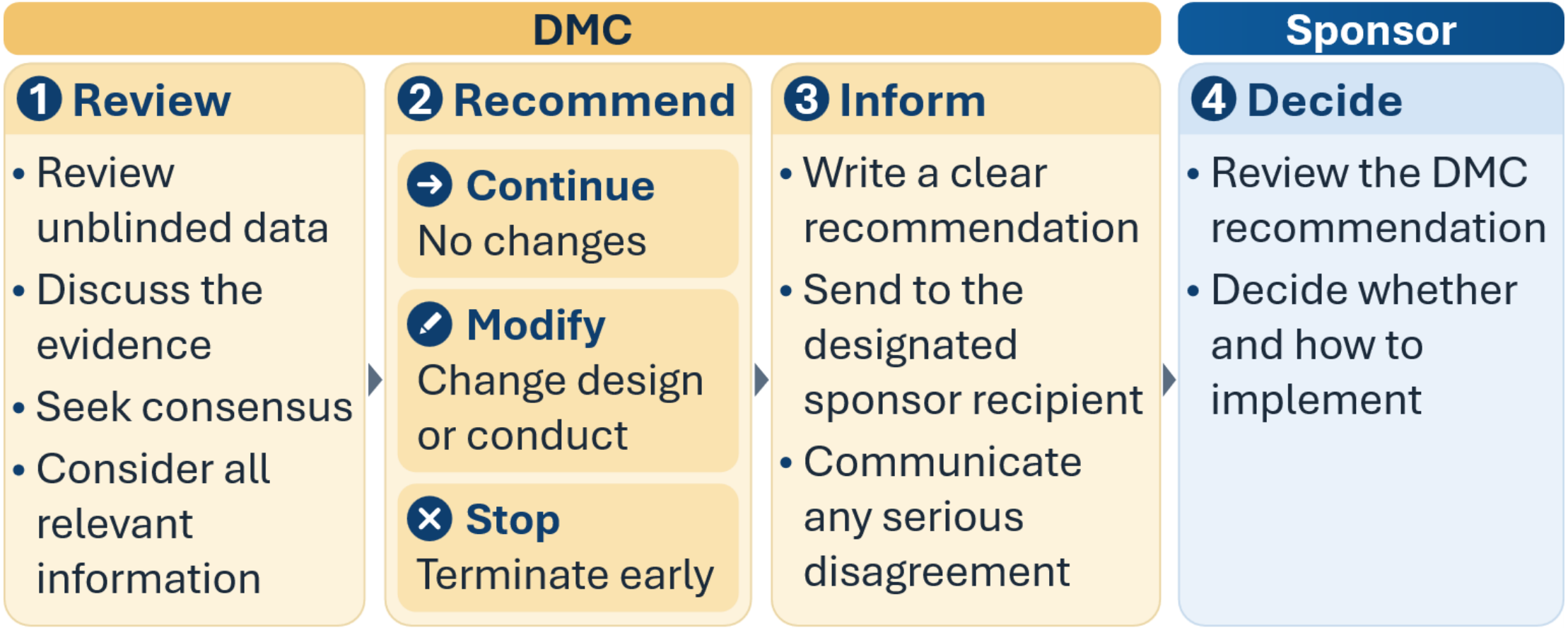
Open vs closed DMC session



Open vs closed DMC reports

	Open report	Access	Closed report
Purpose	Provide trial context without unblinding		Independent, confidential review to recommend
Readers	All open-session attendees		DMC and independent statistician(s) (e.g. SDAC)
Content	• Trial progress & conduct • Non-confidential updates • Blinded, pooled data		• Unblinded safety/efficacy • Pre-specified analyses • DMC-requested analyses
Author	Clinical trial team (sponsor)		Independent statistician / SDAC

DMC recommends, Sponsor decides



DMC recommendation form

- What it captures
  - DMC's recommendation to continue, modify, or stop the trial, communicated clearly to the designated sponsor recipient
  - Sufficient but limited information for the Sponsor to decide on the recommendation

- What should remain outside it
  - Full confidential deliberation
  - (Unnecessary) unblinding information

- Some DMCs prefer a template for providing the recommendation
- Signed by DMC Chair

Summary: DMC's key responsibilities

Study planning

- Review protocol, investigator brochure (IB), etc.
- Review DMC documents (charter, SAP, report formats (TFLs), etc)

Study start-up

- Organizational/kick-off meeting with sponsor
- Finalize charter and monitoring guidelines

Study conduct

- Participate in regular data review meetings
- Monitor unblinded safety and efficacy data, trial integrity, and data quality
- Recommend continuation, modification, or termination

Study close-out

- Final review of data and study results

Diving deeper into specific DMC operations

- Communication between DMC and sponsor
- DMC meeting minutes
- Data access, cleanliness, and timeliness
- Role of the SDAC

Communication between DMC and sponsor

- Direct interactions between DMC members and the Sponsor have risk of unblinding or introducing potential bias
  - Interactions need to be carefully documented and managed
  - Interactions between unblinded DMC members and the trial team should be limited to avoid unblinding and bias.
- Communication policies between DMC and Sponsor are detailed in the DMC Charter
- Sponsor might appoint (senior) member(s) of leadership team as point of contact for safety concerns, DMC recommendation, and any other issues that might require escalation

Diving deeper into specific DMC operations

- Communication between DMC and sponsor
- DMC meeting minutes
- Data access, cleanliness, and timeliness
- Role of the SDAC

Minutes of closed and open session

Record	What it should capture	What should remain outside it	Who takes the minutes?
Open minutes	Content of open session including trial-conduct and operational information, DMC questions and responses	Any unblinding information	Sponsor representative
Closed minutes	Key points influencing the recommendation, confidential analysis requests, DMC-internal follow-up actions	TFL-by-TFL copy of the closed report, verbatim account of everything said	DMC Chair responsible, but might delegate

Closed minutes preserve the reasoning behind DMC decisions

Why they matter

Preserve the rationale
Capture key discussion points that influenced the recommendation.

Maintain continuity
Record outstanding issues, resolutions, and responses to earlier recommendations.

Document unresolved disagreement
Record persistent differences of opinion; do not imply unanimity.

Protect confidential deliberations
Keep the detailed record separate from sponsor-facing communication.

Operational aspects

Keep the record focused
Use neutral wording and minimal attribution; reference reports rather than copy them.

Finalize promptly
Suggested practice: draft within ~5 business days; allow ~5 more for review.

Control access and archiving
Keep closed minutes confidential during the trial; use separate records per protocol in a platform DMC.

Provided to sponsor after end of trial
Will be sumitted to HAs

Diving deeper into specific DMC operations

- Communication between DMC and sponsor
- DMC meeting minutes
- Data access, cleanliness, and timeliness
- Role of the SDAC

Data access, cleanliness, and timeliness

- DMCs needs to have access to all data including efficacy data
  - This also applies to ‘safety DMCs’
  - More on this tomorrow…
- Cleanliness of the data in the DMC reports must be balanced with its currency
  - Not necessary for the Sponsor to send DMC fully clean data at the expense of timely information
  - Important to explain to the DMC any limitations of the DMC report in terms to cleanliness and currency
- DMC members should weigh in on cleanliness and currency

Diving deeper into specific DMC operations

- Communication between DMC and sponsor
- DMC meeting minutes
- Data access, cleanliness, and timeliness
- Role of the SDAC

Role of the SDAC (I)

- For many DMC analyses a separate independent group carries out the unblinded (interim) analyses
  - Especially recommended for studies with a registrational intent
- SDAC: Statistical data analysis center
  - Independent academic department or CRO
  - SDAC used as an umbrella term for the independent group who carries out the unblinded interim analyses
- Independent Statistician (IS) is a member of the SDAC

Role of the SDAC (II)

- Independence of SDAC from the Sponsor essential to ensuring study integrity and scientific validity of the study
  - Act as a gatekeeper for the DMC
  - Allows for neutral evaluations, safeguarding against potential partiality from the Sponsor
- SDAC must have access to all unblinded efficacy and safety data, without needing sponsor permission to fulfill DMC requests
- SDAC statistician should have deep knowledge of the specific study to deliver reliable and interpretable insights

DMCs in non-standard settings

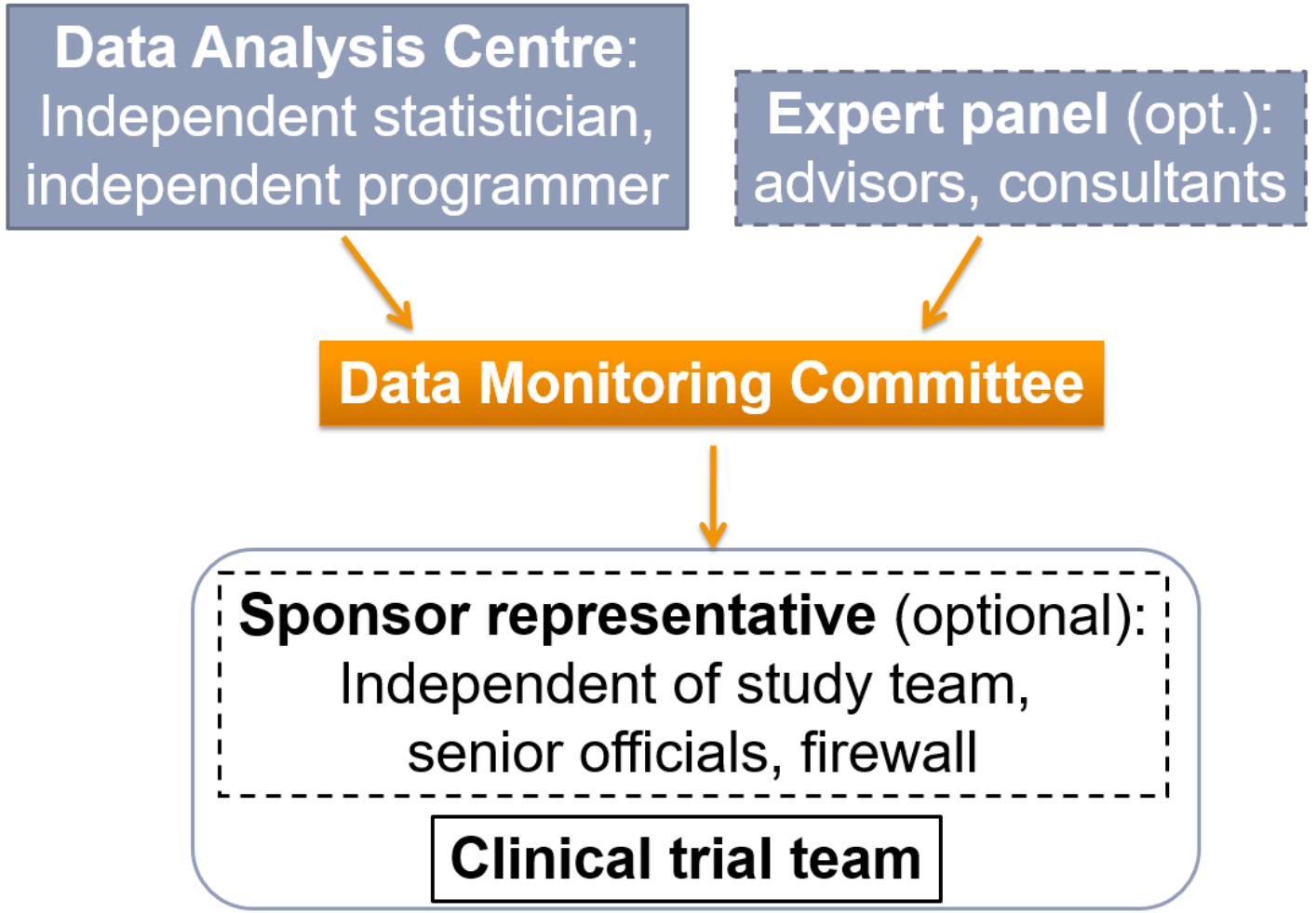
Tim Friede & Marisa Bacchi

Outline

- DMCs in adaptive designs
- Programme level DMC
- Open label trials
- Accelerated approval

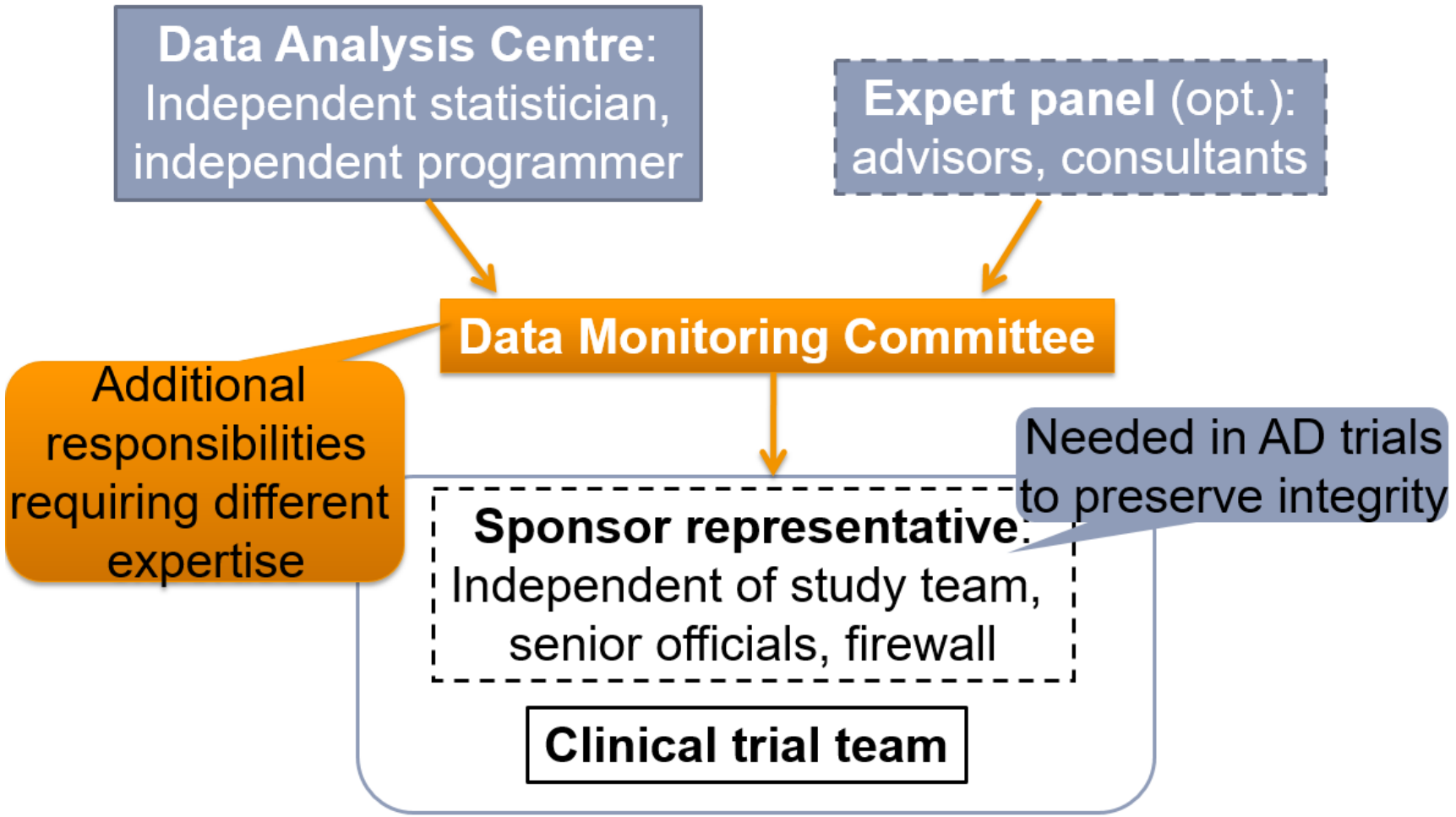
Standard information flow

- The information flow is prospectively specified in the DMC charter



Information flow in adaptive designs

- **Sponsor representative** (also known as sponsor committee) required in clinical trials with adaptive design (unless only blinded adaptations)
- **Additional responsibilities** might require additional expertise on the DMC



Case study:
Adaptive dose-finding study in a rare kidney disease

Received: 26 January 2023 | Revised: 30 June 2023 | Accepted: 3 August 2023
DOI: 10.1002/pst.2335

MAIN PAPER

WILEY

Interim decision making in seamless trial designs: An application in an adaptive dose-finding study in a rare kidney disease

Olympia Papachristofi[1] | Björn Bornkamp[1] | Melanie Wright[1] | Tim Friede[2]

[1]Clinical Development and Analytics, Novartis Pharma AG, Basel, Switzerland
[2]Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

Correspondence
Olympia Papachristofi, Novartis Pharma AG, CH-4002 Basel, Switzerland.
Email: olympia.papachristofi@novartis.com

Funding information
Novartis Pharma AG

Abstract

Adaptive seamless trial designs, combining the learning and confirming cycles of drug development in a single trial, have gained popularity in recent years. Adaptations may include dose selection, sample size re-estimation and enrichment of the study population. Despite methodological advances and recognition of the potential efficiency gains such designs offer, their implementation, including how to enable efficient decision making on the adaptations in interim analyzes, remains a key challenge in their adoption.

This manuscript uses a case study of an adaptive seamless proof-of-concept (Phase 2a)/dose-finding (Phase 2b) to showcase potential adaptive features that can be implemented in trial designs at earlier development stages and the role of simulations in assessing the design operating characteristics and specifying the decision rules for the adaptations. It further outlines the elements needed to support successful interim analysis decision making on the adaptations while safeguarding study integrity, including the role of different stakeholders, interactive simulation-based tools to facilitate decision making and operational aspects requiring preplanning. The benefits of the adaptive Phase 2a/2b design chosen compared to following the traditional two separate studies (2a and 2b) paradigm are discussed.

With careful planning and appreciation of their complexity and components needed for their implementation, seamless adaptive designs have the potential to yield significant savings both in terms of time and resources.

KEYWORDS

adaptive designs, data monitoring committee, dose-finding, IgA nephropathy, proof-of-concept, simulations

- Adaptive seamless proof-of-concept (Phase 2a)/dose-finding (Phase 2b) trial
- Simulations to assess design operating characteristics and to specify decision rules for the adaptations
- Interactive simulation-based tools to facilitate decision making

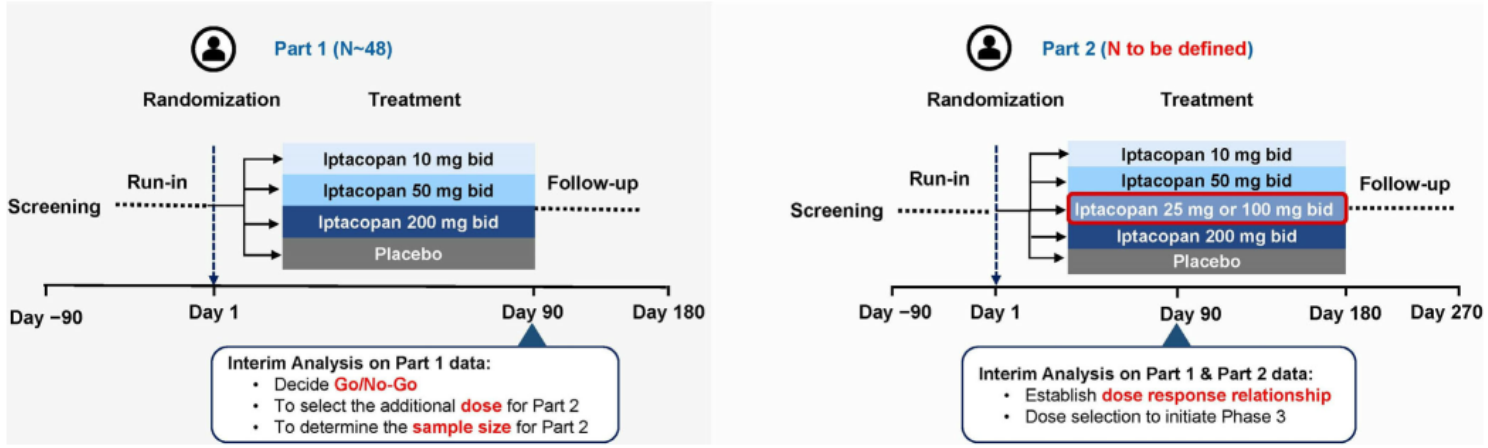


FIGURE 1 The Phase 2a/2b study design. Part 1 refers to the Phase 2a part of the study and Part 2 to the Phase 2b part. b.i.d., twice-daily.

Case study: The role of the DMC

- Investigators, patients and the Sponsor team working on the study remained blinded to the individual treatment codes until final database lock
- Independent DMC established that, besides its “usual” role of conducting periodic safety reviews, also assumed the role of recommending design adaptations
- Three members, two clinicians and one statistician
- All members experienced trialists, and statistician adaptive design methodology expert
- Study protocol described DMC role at a high-level whereas dedicated DMC Charter detailed DMC activities and adaptation rules

Case study: Supporting the DMC in the interim decision making with a simulation-based interactive tool

- Face-to-face **DMC meeting** organized 9 months ahead of interim analysis (IA)
- **Training on adaptations using interactive simulation based RShiny app**
  - useful in engaging with nonstatistical members
  - illustrates variability and consequences for decisions
  - exchanges between the DMC and Sponsor on hypothetical decisions

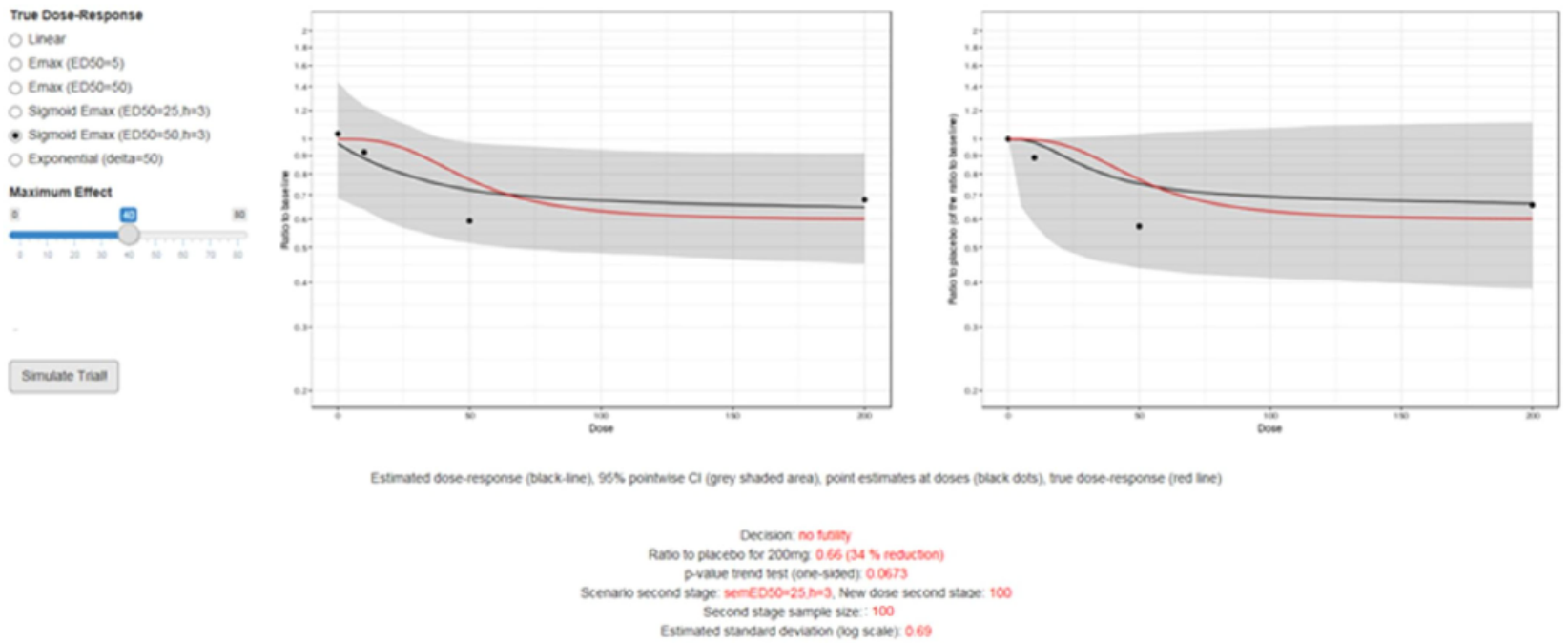


FIGURE 6 Snapshot from the RShiny app used with the DMC. The simulation parameters (i.e., true dose response model and true maximum treatment effect) could be selected at the top left. The plot on the left illustrated dose–response for the ratio to baseline in UPCR and the plot on the right the placebo adjusted dose–response for the ratio to baseline in UPCR. The red line indicates the simulated true dose–response curve and the black line the estimated dose–response curve i.e., the Bayesian sigmoid Emax model (posterior median) with 95% Bayesian credible intervals (grey shaded area). The black dots correspond to the MMRM estimates (i.e., least squares means). DMC, data monitoring committee; MMRM, mixed models for repeated measures; UPCR, urine protein to creatinine ratio.

Outline

- DMCs in adaptive designs
- Programme level DMC
- Open label trials
- Accelerated approval

Program-level DMCs highlighted as a significant change in FDA draft guidance

Significant changes in DMC structure and practice since the 2006 guidance was issued include:

- The increased use of DMCs in trials (Califf et al. 2012) of modest size as reflected in the clinical trials data bank housed at ClinicalTrials.gov[6]
- A trend for DMC charters to become longer and more detailed
- An increased use of DMCs to implement certain adaptive clinical trial designs[7]
- An increased use of some DMCs to oversee an entire clinical development program rather than a single clinical trial[8]
- The potential for expansion of functions of a DMC; for example, for review of aggregate[9] data for safety reporting for trials under an investigational new drug application (IND)
- An increased globalization of medical product development and use of multiregional trials with DMCs[10]

[8] See Clinical Trials Transformation Initiative (CTTI) Overview: Data Monitoring Committees, available at https://ctti-clinicaltrials.org/our-work/ethics-and-human-research-protection/data-monitoring-committees/.

- Summary of changes from previous version (2006) included in draft guidance (2024)
- https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-data-monitoring-committees-clinical-trials

Programme level DMC

- DMC monitoring all (or at least several) clinical trials within a development programme
- Motivation for programme level DMC
  - Rare adverse events can only be discovered in larger samples (i.e. by combining trials)
  - Carry over experience gained in phase II into phase III (i.e. create continuity), in particular with regard to management of specific side effects
- Possibly larger than the usual trial level DMC
  - for example by including a pharmaco-epidemiology expert or experts relating to the specific settings of the individual studies

Integrating data from several trials

Standard practice in (IPD) meta-analysis to be followed including …

- Presenting data by trial
  - allowing assessment of between-trial heterogeneity
- Stratifying combined analyses by trial
  - to deal with between-trial heterogeneity in (control) event rates
- Considering between-trial heterogeneity in treatment effects
  - With “study twins” (trials with the same protocol) a common-effect (fixed-effect) model might be appropriate
  - When common-effect model appears inappropriate (e.g. due to difference in doses or dose frequencies, random-effects model should be used. With few trials (typical in programme level DMCs) a Bayesian framework advantageous (Unkel et al, 2019)

Outline

- DMCs in adaptive designs
- Programme level DMC
- Open label trials
- Accelerated approval

Open label trial – Is a DMC needed?



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Transcatheter Valve Repair in Heart Failure with Moderate to Severe Mitral Regurgitation

S.D. Anker, T. Friede, R.-S. von Bardeleben, J. Butler, M.-S. Khan, M. Diek, J. Heinrich, M. Geyer, M. Placzek, R. Ferrari, W.T. Abraham, O. Alfieri, A. Auricchio, A. Bayes-Genis, J.G.F. Cleland, G. Filippatos, F. Gustafsson, W. Haverkamp, M. Kelm, K.-H. Kuck, U. Landmesser, A.P. Maggioni, M. Metra, V. Ninios, M.C. Petrie, T. Rassaf, F. Ruschitzka, U. Schäfer, P.C. Schulze, K. Spargias, A. Vahanian, J.L. Zamorano, A. Zeiher, M. Karakas, F. Koehler, M. Lainscak, A. Öner, N. Mezilis, E.K. Theofilogiannakos, I. Ninios, M. Chrissoheris, P. Kourkoveli, K. Papadopoulos, G. Smolka, W. Wojakowski, K. Reczuch, F.J. Pinto, Ł. Wiewiórka, Z. Kalarus, M. Adamo, E. Santiago-Vacas, T.F. Ruf, M. Gross, J. Tongers, G. Hasenfuss, W. Schillinger, and P. Ponikowski, for the RESHAPE-HF2 Investigators*

ABSTRACT

BACKGROUND
Whether transcatheter mitral-valve repair improves outcomes in patients with heart failure and functional mitral regurgitation is uncertain.

METHODS
We conducted a randomized, controlled trial involving patients with heart failure and moderate to severe functional mitral regurgitation from 30 sites in nine countries. The patients were assigned in a 1:1 ratio to either transcatheter mitral-valve repair and guideline-recommended medical therapy (device group) or medical therapy alone (control group). The three primary end points were the rate of the composite of first or recurrent hospitalization for heart failure or cardiovascular death during 24 months; the rate of first or recurrent hospitalization for heart failure during 24 months; and the change from baseline to 12 months in the score on the Kansas City Cardiomyopathy Questionnaire–Overall Summary (KCCQ-OS; scores range from 0 to 100, with higher scores indicating better health status).

RESULTS
A total of 505 patients underwent randomization: 250 were assigned to the device group and 255 to the control group. At 24 months, the rate of first or recurrent hospitalization for heart failure or cardiovascular death was 37.0 events per 100 patient-years in the device group and 58.9 events per 100 patient-years in the control group (rate ratio, 0.64; 95% confidence interval [CI], 0.48 to 0.85; P=0.002). The rate of first or recurrent hospitalization for heart failure was 26.9 events per 100 patient-years in the device group and 46.6 events per 100 patient-years in the control group (rate ratio, 0.59; 95% CI, 0.42 to 0.82; P=0.002). The KCCQ-OS score increased by a mean (±SD) of 21.6±26.9 points in the device group and 8.0±24.5 points in the control group (mean difference, 10.9 points; 95% CI, 6.8 to 15.0; P<0.001). Device-specific safety events occurred in 4 patients (1.6%).

CONCLUSIONS
Among patients with heart failure with moderate to severe functional mitral regurgitation who received medical therapy, the addition of transcatheter mitral-valve repair led to a lower rate of first or recurrent hospitalization for heart failure or cardiovascular death and a lower rate of first or recurrent hospitalization for heart failure at 24 months and better health status at 12 months than medical therapy alone. (Funded by Abbott Laboratories; RESHAPE-HF2 ClinicalTrials.gov number, NCT02444338.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Dr. Anker can be contacted at s.anker@cachexia.de or at the Department of Cardiology (CVK), Charité Universitätsmedizin, D-13353 Berlin, Germany.

*A list of the RESHAPE-HF2 investigators is provided in the Supplementary Appendix, available at NEJM.org.

This article was published on August 31, 2024, and last updated on November 14, 2024, at NEJM.org.

N Engl J Med 2024;391:1799-809.
DOI: 10.1056/NEJMoa2314328

Copyright © 2024 Massachusetts Medical Society.

CME

The New England Journal of Medicine is produced by NEJM Group, a division of the Massachusetts Medical Society.
Downloaded from nejm.org by TIM FRIEDE on November 14, 2024. For personal use only.
No other uses without permission. Copyright © 2024 Massachusetts Medical Society. All rights reserved.

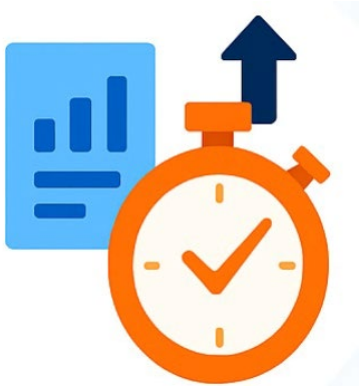
- **Open-label, randomized, confirmatory/pivotal trials should be treated similarly as fully blinded trials**
- Assess the need for a DMC – if there is a need for assessment of safety, futility, and efficacy for ongoing trial
- Access to summary of information by treatment arm should be limited to independent DMCs to maintain study integrity
- **Example: RESHAPE-HF2**, a randomized controlled device trial with blinded outcome assessment (PROBE design)
  - During the course of the study, the steering committee and trial team had only access to "blinded" data pooled across treatment arms
  - An independent data and safety monitoring committee evaluated patient safety during the trial

Outline

- DMCs in adaptive designs
- Programme level DMC
- Open label trials
- Accelerated approval

DMCs in clinical trials for accelerated approval

- Interim Analysis based on a surrogate efficacy endpoint supportive of accelerated (but not full!) approval → study to continue blinded to final analysis on confirmatory endpoint
- Submission decision may not only be gated by a "successful" primary analysis:
  - Totality of efficacy data, including confirmatory endpoint at final analysis may be needed.
  - "Success" on other efficacy endpoints may be needed depending on the competitive landscape
- Submission decision needed fast to maintain "gains" of AA approach & avoid negative perception with potential impact on adherence
- Sponsor access to unblinded data required for submission and interactions with regulatory authorities.



DMCs in clinical trials for accelerated approval

Two (at least) operational models - early consultation with regulatory authorities to establish agreement on the operational model proposed highly encouraged

- Option 1
  - SDAC conducts primary analysis → DMC makes submission recommendation → limited sponsor team, not involved in trial conduct, unblinded for submission.
  - May be preferable for smaller companies where blinded/unblinded teams are not feasible and operational submission-enabling activities are outsourced (e.g. programming of datasets, statistical analysis and reporting, regulatory writing)
- Option 2
  - Access to FIR restricted to independent programmer/statistician and limited senior management at sponsor to enable "go" decision
  - May be preferable for large-scale companies where separate blinded/unblinded teams can be established and operational submission-enabling activities are done inhouse

Documentation of measures/firewalls in place to maintain blind needed irrespective of model chosen.

Title

- Text

Overview of statistical methods for monitoring

Tim Friede

Introduction

- **Statistical monitoring of estimators**
  - Treatment effects (differences between treatment arms)
  - Nuisance parameters (e.g. variances, overall event rate)
- **Sequential analyses**
  - Strictly sequential (data look after every new observation)
  - Group sequential (small number of data looks during the course of the trial)
- Here we focus on **group sequential monitoring of treatment effects**
  - Nuisance parameters typically monitored by trial team in a blinded fashion (without knowledge of treatment group allocations) and therefore not of concern to DMC
  - Strictly sequential designs often not feasible

Fixed (non-sequential) designs

- In the non-sequential case, the sample size, n_f, is fixed and one comes to a decision after observing the complete sample.
- Consider independent and normally distributed observations $X_1, \ldots, X_{n_f}$ with unknown mean and known variance σ^2.
- Let the null hypothesis to be tested at significance level α, be given by $H_0: \mu = \mu_0$.
- The test decision is based on the statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n_f}$$

- Let H_0 be tested against the two-sided alternative, which is given by $H_1: \mu \neq \mu_0$.
- H_0 is rejected if the observed absolute value $|z|$ of the test statistic Z exceeds the critical value $u = \phi^{-1}(1 - \alpha/2)$ where ϕ^{-1} denotes the inverse standard normal cumulative distribution function (cdf).

Group sequential designs

- Repeated analyses (so-called data looks) on accumulating data
- At each data look: decision to continue the study, to stop for efficacy or to stop for futility

Treatment effect β (e.g. mean difference, log odds ratio, log hazard ratio)

$$\underbrace{X_{11},\ldots,X_{1n_1}}_{n_1 \text{ observations}},\underbrace{X_{21},\ldots,X_{2n_2}}_{n_2 \text{ observations}},\ \ldots\ ,\underbrace{X_{K1},\ldots,X_{Kn_K}}_{n_K \text{ observations}}$$

Stage 1 $\quad Z_1 = \dfrac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$

Stage 2 $\quad Z_2 = \dfrac{\hat{\beta}_2}{SE(\hat{\beta}_2)}$

Stage K $\quad Z_K = \dfrac{\hat{\beta}_K}{SE(\hat{\beta}_K)}$

Group sequential designs

Treatment effect β (e.g. mean difference, log odds ratio, log hazard ratio)

$$\underbrace{X_{11}, \ldots, X_{1n_1}}_{n_1 \text{ observations}}, \underbrace{X_{21}, \ldots, X_{2n_2}}_{n_2 \text{ observations}}, \ldots, \underbrace{X_{K1}, \ldots, X_{Kn_K}}_{n_K \text{ observations}}$$

Stage 1 $Z_1 = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$

Stage 2 $Z_2 = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}$

Stage K $Z_K = \frac{\hat{\beta}_K}{SE(\hat{\beta}_K)}$

- Usually the same test statistics as in the fixed designs
- Critical values (so-called stopping boundaries) adjusted to control the overall (familiywise) type I error rate

Repeated testing

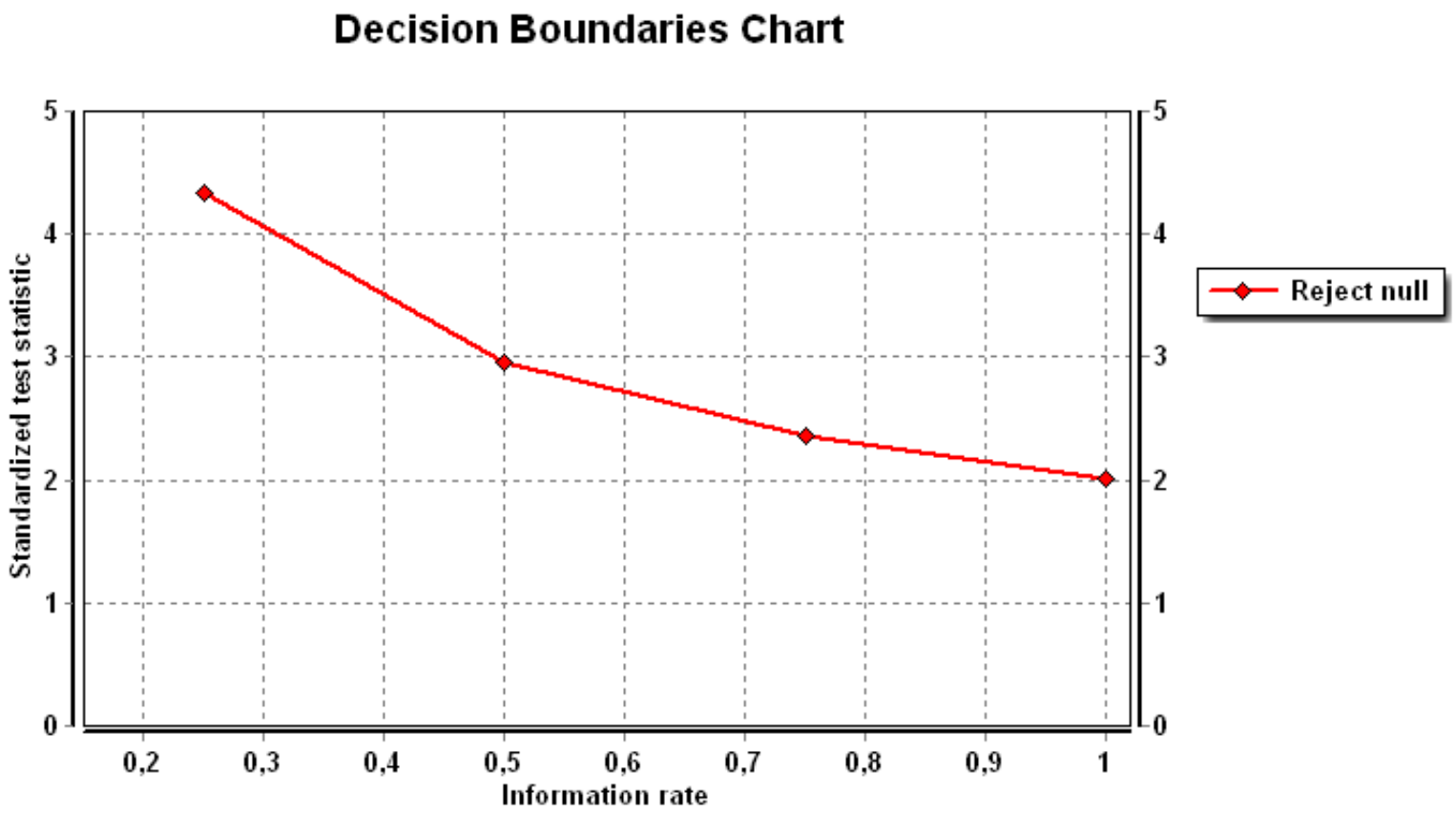
Table 1.1 Type I error rate when using the unadjusted critical value $u = \Phi^{-1}(1 - \alpha/2)$ at each stage of the test procedure in a two-sided test design

	α	0.001	0.01	0.05
K	u	3.291	2.576	1.960
1		0.00100	0.01000	0.05000
2		0.00186	0.01766	0.08312
3		0.00257	0.02366	0.10726
4		0.00319	0.02858	0.12617
5		0.00372	0.03274	0.14169
10		0.00569	0.04738	0.19336
15		0.00705	0.05692	0.22509
20		0.00808	0.06403	0.24791
25		0.00892	0.06971	0.26567
30		0.00963	0.07444	0.28016
35		0.01025	0.07849	0.29238
40		0.01079	0.08204	0.30293
45		0.01128	0.08519	0.31220
50		0.01172	0.08803	0.32045

The stage sizes are assumed to be equal

- Table 1.1 from Wassmer & Brannath (2016)
- Repeated tests are carried out without adjustment of the critical values
- Inflation of type I error rate
- Therefore, critical values need to be adjusted to control the error rate across repeated tests

Group sequential designs: Stopping boundaries



K = 4; alpha = 0,025, one-sided, O'Brien and Fleming type alpha spending

- **Example**: Plot displaying decision boundaries (produced with ADDPLAN)
- Here only **early stopping for efficacy**; non-binding futility boundaries could be used in addition

Why group sequential designs?

- **Ethical reasons**, e.g. long-running studies, vulnerable patients
- So far **internal pilot studies** considered. These **mitigate uncertainty in nuisance parameters**, but assume a smallest clinically relevant effect
- Group-sequential designs **mitigate uncertainty with regard to treatment effect**
- For instance, expected effect much larger than smallest clinically relevant effect
- **Assumption on effect size has big impact on sample size**, often a factor of f in effect size results in a factor of $1/f^2$ in sample size

$$\mathrm{n} = \frac{(k+1)^2}{k} \frac{(\Phi^{-1}(\alpha) + \Phi^{-1}(\beta))^2}{\Delta^{*2}} \sigma^2$$

Group sequential designs: Stopping boundaries

For example, consider H_0: $\mu = \mu_0$ *vs* H_1: $\mu > \mu_0$

$$\underbrace{X_{11}, \ldots, X_{1n_1}}_{n_1 \text{ observations}}, \underbrace{X_{21}, \ldots, X_{2n_2}}_{n_2 \text{ observations}}, \ldots, \underbrace{X_{K1}, \ldots, X_{Kn_K}}_{n_K \text{ observations}}$$

Stage 1 $\underbrace{\qquad}_{Z_1 = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}}$

Stage 2 $\underbrace{\qquad\qquad}_{Z_2 = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}}$

Stage K $\underbrace{\qquad\qquad\qquad}_{Z_K = \frac{\hat{\beta}_K}{SE(\hat{\beta}_K)}}$

After stage $k = 1, \ldots, K-1$

- if $Z_k \geq c_k$ stop, reject H_0
- otherwise continue to group $k+1$,

after stage K

- if $Z_K \geq c_K$ stop, reject H_0
- otherwise stop, accept H_0.

How to choose critical values?

These critical values will help us control the Type I error rate.

Test statistics and canonical distribution

- **Treatment effect** β (e.g. mean difference, log odds ratio, log hazard ratio)
- **Z-statistic** at data look $k = 1, \dots, K$: $Z_k = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}$
- **Information level** at data look k: $\mathcal{I}_k = \frac{1}{SE(\hat{\beta}_k)^2}$
- **Information fraction** at look k: $\mathrm{w_k} = \frac{\mathcal{I}_k}{\mathcal{I}_{Max}}$
- **Canonical joint distribution**
  - Vector of test statistics $(Z_1, \dots, Z_K)$ follows a **multivariate normal distribution** with expectation $E[Z_k] = \beta\mathcal{I}_{\mathrm{k}}$ $(k = 1, \dots, K)$ and covariance $Cov[Z_{k_1}, Z_{k_2}] = \sqrt{\mathcal{I}_{\mathrm{k_1}}/\mathcal{I}_{\mathrm{k_2}}}$ $(1 \leq k_1 \leq k_2 \leq K)$

Type I error rate

			Test decision	
			Not to reject H_0	Reject H_0
Truth (unknown)	H_0	H_0 true	correct	Type I error **alpha**
	H_1	H_0 false	Type II error **beta**	correct **Power:1-beta**

After stage $k = 1, \ldots, K-1$

- if $Z_k \geq c_k$ stop, reject H_0
- otherwise continue to group $k+1$,

after stage K

- if $Z_K \geq c_K$ stop, reject H_0
- otherwise stop, accept H_0.

- **Continuation region** at stage k denoted by C_k, the complement of the continuation region is $\bar{C}_k$. Here we assume that the **rejection region** R_k is equal to the complement of the continuation region $\bar{C}_k$.
- Study continues if $Z_k \in C_k$, for $k = 1, \ldots, K-1$
- Null hypothesis H_0 rejected at stage k if $Z_k \in \bar{C}_k$

Type I error rate

		Test decision	
		Not to reject H_0	Reject H_0
Truth (unknown)	H_0 H_0 true	correct	Type I error **alpha**
	H_1 H_0 false	Type II error **beta**	correct **Power:1-beta**

After stage $k = 1, \ldots, K-1$
- if $Z_k \geq c_k$ stop, reject H_0
- otherwise continue to group $k+1$,

after stage K
- if $Z_K \geq c_K$ stop, reject H_0
- otherwise stop, accept H_0.

- Study continues if $Z_k \in C_k$, for $k = 1, \ldots, K-1$
- Null hypothesis H_0 rejected at stage k if $Z_k \in \bar{C}_k$
- Design controls the type I error rate at level α if the following holds:

$$P_{H_0}(Z_1 \in \bar{C}_1 \; or \; Z_2 \in \bar{C}_2 \; or \; \cdots \; or \; Z_K \in \bar{C}_K)$$

$$= P_{H_0}\left(\bigcup_{k=1}^{K}\{Z_k \in \bar{C}_k\}\right) = 1 - P_{H_0}\left(\bigcap_{k=1}^{K}\{Z_k \in C_k\}\right) \leq \alpha$$

Type I error rate

- Design controls the type I error rate at level α if the following holds:

$$P_{H_0}(Z_1 \in \bar{C}_1 \; or \; Z_2 \in \bar{C}_2 \; or \; \cdots \; or \; Z_K \in \bar{C}_K)$$

$$= P_{H_0}\left(\bigcup_{k=1}^{K}\{Z_k \in \bar{C}_k\}\right) = 1 - P_{H_0}\left(\bigcap_{k=1}^{K}\{Z_k \in C_k\}\right) \leq \alpha$$

- Denoting the density of the **multivariate normal distribution** by $\boldsymbol{f(z_1, \ldots, z_K)}$ with expectation $\boldsymbol{E[Z_k] = 0}$ $(k = 1, \ldots, K)$ and covariance $\boldsymbol{Cov[Z_{k_1}, Z_{k_2}] = \sqrt{\mathcal{I}_{k_1}/\mathcal{I}_{k_2}}}$ $(1 \leq k_1 \leq k_2 \leq K)$ the type I error rate can be computed as

$$1 - \int_{C_K}^{\square} \cdots \int_{C_1}^{\square} f(z_1, \ldots, z_k)\, dz_1 \cdots dz_K$$

- Technically, this is done by use of the so-called **recursive integration formula** (Armitage et al, 1969)

Type I error rate

- Design controls the type I error rate at level α if the following holds:

$$P_{H_0}(Z_1 \in \bar{C}_1 \; or \; Z_2 \in \bar{C}_2 \; or \; \cdots \; or \; Z_K \in \bar{C}_K)$$

$$= P_{H_0}\left(\bigcup_{k=1}^{K}\{Z_k \in \bar{C}_k\}\right) = 1 - P_{H_0}\left(\bigcap_{k=1}^{K}\{Z_k \in C_k\}\right) \leq \alpha$$

- An alternative expression :

$$P_{H_0}(Z_1 \in \bar{C}_1 \; or \; Z_2 \in \bar{C}_2 \; or \; \cdots \; or \; Z_K \in \bar{C}_K)$$
$$= P_{H_0}(Z_1 \in \bar{C}_1) + P_{H_0}(Z_1 \in C_1, Z_2 \in \bar{C}_2) + \cdots + P_{H_0}(Z_1 \in C_1, Z_2 \in C_2, \ldots, Z_K \in \bar{C}_K)$$

Could we allocate the overall Type I error rate to each stage?

Stopping boundaries

- **Error spending function**
  - Non-decreasing function $f: [0, \infty) \rightarrow [0, \alpha]$ with $f(0) = 0$ and $f(t) = \alpha$ for $t \geq 1$
- **Type I error rate π_k allocated to look $k = 1, \ldots, K$:**
  - $k = 1 : f\left(\frac{\mathcal{I}_1}{\mathcal{I}_{max}}\right)$
  - $k = 2, \ldots, K: f\left(\frac{\mathcal{I}_k}{\mathcal{I}_{max}}\right) - f(\frac{\mathcal{I}_{k-1}}{\mathcal{I}_{max}})$
- **Stopping boundaries** (critical values) determined by plugging in estimate for information level at look k in formula above and by estimating covariance matrix of the canonical distribution.

Allocate Type I error rate to each stage

Stage 1	$P_{H_0}(Z_1 \in \bar{C}_1)$	$= f\left(\frac{\mathcal{I}_1}{\mathcal{I}_{max}}\right)$
Stage 2	$P_{H_0}(Z_1 \in C_1, Z_2 \in \bar{C}_2)$	$= f\left(\frac{\mathcal{I}_2}{\mathcal{I}_{max}}\right) - f\left(\frac{\mathcal{I}_1}{\mathcal{I}_{max}}\right)$
Stage K	$P_{H_0}(Z_1 \in C_1, Z_2 \in C_2, \dots, Z_K \in \bar{C}_K)$	$= \alpha - f\left(\frac{\mathcal{I}_{K-1}}{\mathcal{I}_{max}}\right)$

α

$$P_{H_0}(Z_1 \in \bar{C}_1 \text{ or } Z_2 \in \bar{C}_2 \text{ or } \cdots \text{ or } Z_K \in \bar{C}_K)$$
$$= P_{H_0}(Z_1 \in \bar{C}_1) + P_{H_0}(Z_1 \in C_1, Z_2 \in \bar{C}_2) + \cdots + P_{H_0}(Z_1 \in C_1, Z_2 \in C_2, \dots, Z_K \in \bar{C}_K)$$

Numerical example

- Numerical example of a group sequential plan:
  - α=0.025 (one-sided), K=4 equally spaced data looks
  - Spending functions: O'Brien-Fleming, Pocock

	O'Brien-Fleming		Pocock	
Data look	z_i	Nominal α	z_i	Nominal α
1	4.049	<0.001	2.361	0.009
2	2.863	0.002	2.361	0.009
3	2.338	0.010	2.361	0.009
4	2.024	0.021	2.361	0.009

Numerical example

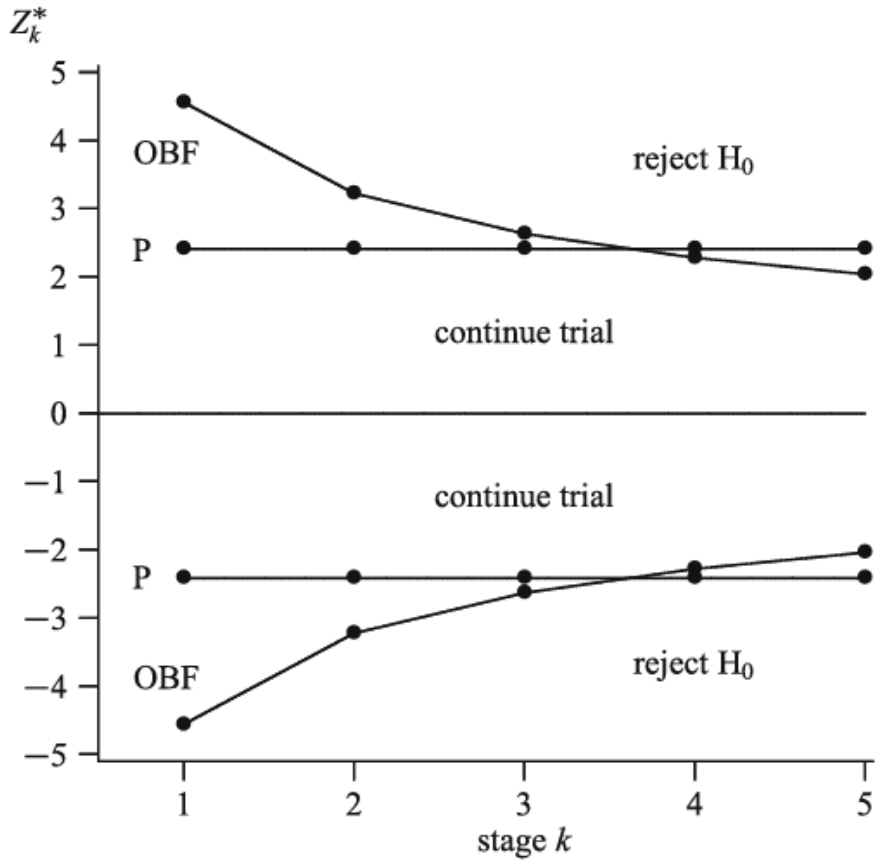


Fig. 2.1 Decision regions for O'Brien and Fleming's (OBF) and Pocock's (P) design for equal stage sizes; $K = 5, \alpha = 0.05$

- Figure 2.1 from Wassmer and Brannath (2016)
- Numerical example of a group sequential plan:
  - α=0.05 (two-sided), K=5 equally spaced data looks
  - Spending functions: O'Brien-Fleming, Pocock

Power

- Continuation and rejection regions at stage k denoted by C_k and R_k, respectively, i.e. study continues if $Z_k \in C_k$ and null hypothesis $H_0: \beta = 0$ rejected at stage k if $Z_k \in R_k$
- For a specific alternative hypothesis $H_1: \beta = \beta^*$ and sample sizes $n_1, \ldots, n_K$ the power is given by

$$P_{H_1}(Z_1 \in R_1) + P_{H_1}(Z_1 \in C_1, Z_2 \in R_2) + \cdots + P_{H_1}(Z_1 \in C_1, \ldots, Z_{K-1} \in C_{K-1}, Z_K \in R_K)$$

- Under the alternative hypothesis $H_1: \beta = \beta^*$ the vector $(Z_1, \ldots, Z_K)$ follows a multivariate normal distribution with expectation $E[Z_k] = \beta^* \mathcal{I}_{\mathrm{k}}$ $(k = 1, \ldots, K)$ and covariance $Cov\left[Z_{k_1}, Z_{k_2}\right] = \sqrt{\mathcal{I}_{\mathrm{k}_1} / \mathcal{I}_{\mathrm{k}_2}}$ $(1 \leq k_1 \leq k_2 \leq K)$

Average sample size

- **Note**, in a group sequential design the sample size is a random variable
- The expected or **average sample size under the alternative** is referred to by ASN_{H_1} and is given by

$$ASN_{H_1} = n_1 + \sum_{k=2}^{K} n_k P\left(\bigcap_{l=1}^{k-1} Z_l \in C_k\right)$$

Power and average sample size

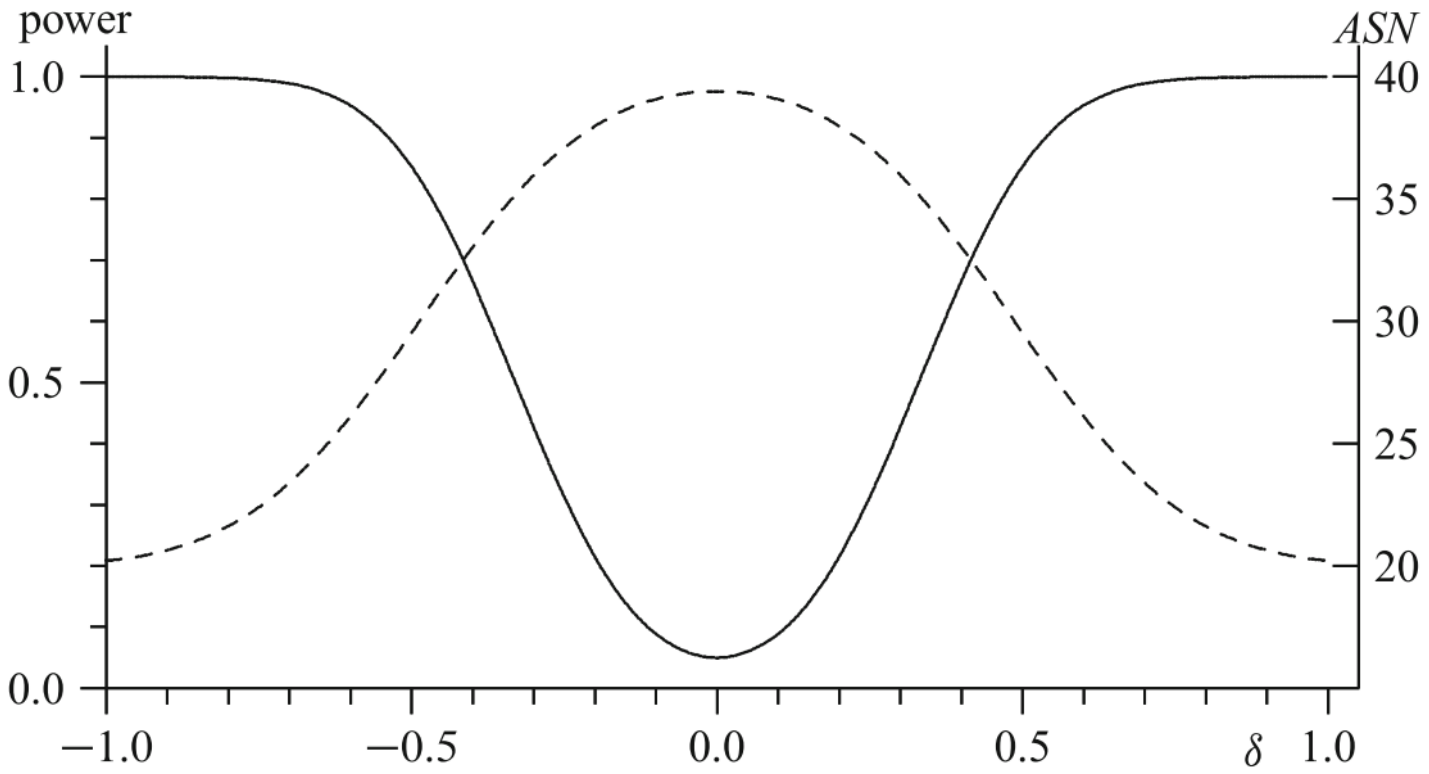


Fig. 1.4 Power (*solid line*) and average sample size (*dashed line*) of a two-stage group sequential test design for testing $H_0 : \mu = \mu_0$ with critical value 2.178 for both stages and continuation region $\mathscr{C}_1^* = (-2.178, 2.178)$. The stage sample sizes are $n_1 = n_2 = 20$

- Figure 1.4 from Wassmer and Brannath (2016)

Power and average sample size

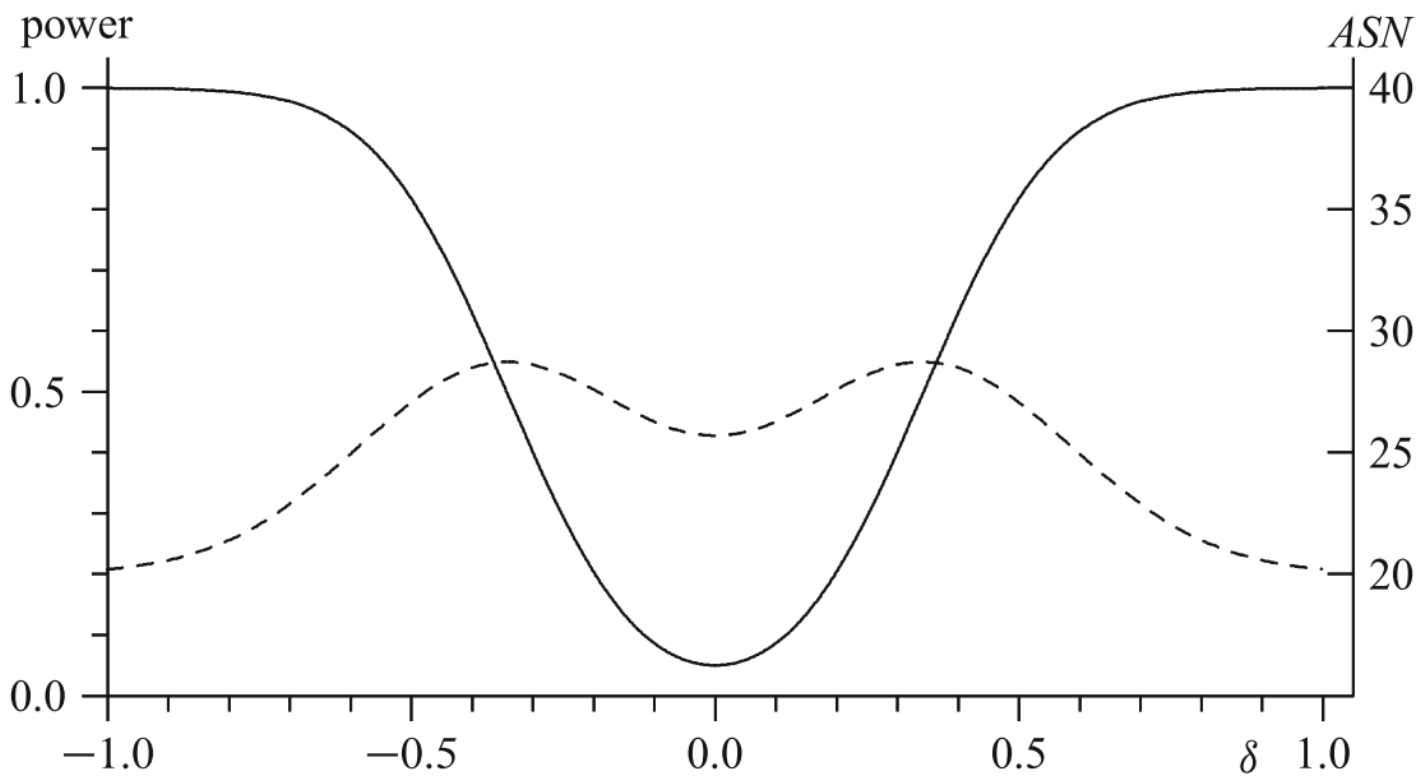


Fig. 1.3 Power (*solid line*) and average sample size (*dashed line*) of a two-stage group sequential test design for testing $H_0 : \mu = \mu_0$ with critical value 2.140 for both stages and continuation region $\mathscr{C}_1^* = (-2.140, -1.0) \cup (1.0, 2.140)$. The stage sample sizes are $n_1 = n_2 = 20$

- Figure 1.3 from Wassmer and Brannath (2016)
- Futility stopping

Some commonly used futility criteria

- **Conditional power**
  - given the treatment effect assumed in the planning
  - given the observed treatment effect
- **Predictive power**
  - Bayesian probability based on the posterior of the treatment effect
- **Under certain assumptions these can be converted one into another**
- **Reference**: Gallo et al (2014) JBS

Some commonly used futility criteria

In this article, CP denotes the conditional power that conditions on the prespecified Δ, and CP(d) denotes the conditional power computed as if the observed estimate is the effect size governing the rest of the trial. For an analysis at information time I, these are well known or easily determined to be of the following forms:

$$CP = \Phi\left(-\frac{1}{\sqrt{1-I}}\left(Iz_\alpha - \sqrt{I}z_I - (1-I)\,z_\beta\right)\right) \quad CP(d) = \Phi\left(-\frac{1}{\sqrt{1-I}}\left(z_\alpha - \frac{z_I}{\sqrt{I}}\right)\right)$$

Lachin (2005) presents a review of conditional power approaches for futility stopping.

Predictive power: This is determined in a Bayesian framework based on a posterior distribution for θ given the interim results. Most commonly, a noninformative prior is updated based on the interim data, but a more general framework allowing an informative prior could certainly be used. With PP denoting the predictive power computed using an initial noninformative prior, this can be expressed as follows:

$$PP = \Phi\left(-\frac{1}{\sqrt{1-I}}\left(\sqrt{I}z_\alpha - z_I\right)\right)$$

- Gallo et al (2014) JBS

Futility criteria

- Table 1 in Gallo et al (2014) JBS

Table 1 Power loss versus stopping under H_0 for futility rules expressed on varying scales (2.5% level, 90% power for Δ, futility look at $I = 0.50$)

Futility rule expression scale					Scheme properties	
Z-score	d/Δ	CP	CP(d)	PP	Power loss	Stop under H_0
No stopping	—	—	—	—	0.0%	0%
0.00	0.00	32%	< 1%	3%	0.2%	50%
0.25	0.11	41%	1%	5%	0.6%	60%
0.50	0.22	51%	4%	11%	1.3%	69%
0.75	0.33	61%	10%	18%	2.7%	77%
1.00	0.44	70%	22%	29%	5.1%	84%

Other outcomes

- With normal outcomes and unknown variance or non-normal endpoints, multivariate normal distribution not exact
- However, standard practice to use **normal approximation**
- **Approximations based on t-distributions** proposed (Pocock, 1977), but not routinely used
- **Time-to-event outcomes** (e.g. survival)
  - Under the null hypothesis (H0: HR=1) variance depends only on the total number of events
  - Event-driven designs common (trial complete when a total number of events is reached): power at a certain alternative guaranteed (but length of trial variable)

Safety monitoring

- **Safety monitoring on different levels**
  - Patient
  - Centre / site
  - Treatment arm (or dose cohort)
  - Trial
- **Possible consequences if safety issues arise** may include …
  - Suspend / stop treatment of individual patients
  - Suspend / stop treatment of patients in certain centres / sites
  - Suspend / close recruitment/ randomization of new patients
  - Suspend / stop dosing of patients
  - Termination of clinical trial

Safety monitoring rules in early clinical trials

- **Safety monitoring rules in early clinical trials** typically based on …
  - Number of dose limiting toxicities (DLT) in a dose cohort
  - Proportion of patients with serious adverse events (SAE) or treatment related adverse events of NCI-CTCAE grade 3 or higher
  - …
- **Bayesian stopping rules**
  - With sparse data (e.g. small dose cohorts), Bayesian framework has often some advantages over frequentist or likelihood frameworks
  - For instance, phase I dose escalation trials with Bayesian logistic regression model (Neuenschwander et al (2008) Statistics in Medicine)

Stopping boundaries for safety

- **Group-sequential plans**
  - well established to monitor efficacy
  - control type I error rate of the trial at level α
  - may also be applied to monitoring harm, however, they generally have low power to detect existing harm, especially early in the trial
- **Alternatively, to increase the probability that the monitoring procedure detects existing harm**, hypothesis test at each look may be performed with a nominal significance level of α.
- **Properties of stopping rules** (such as error rates): to be explored in simulations
- See Mütze & Friede (2020) CCT for an example in COVID-19

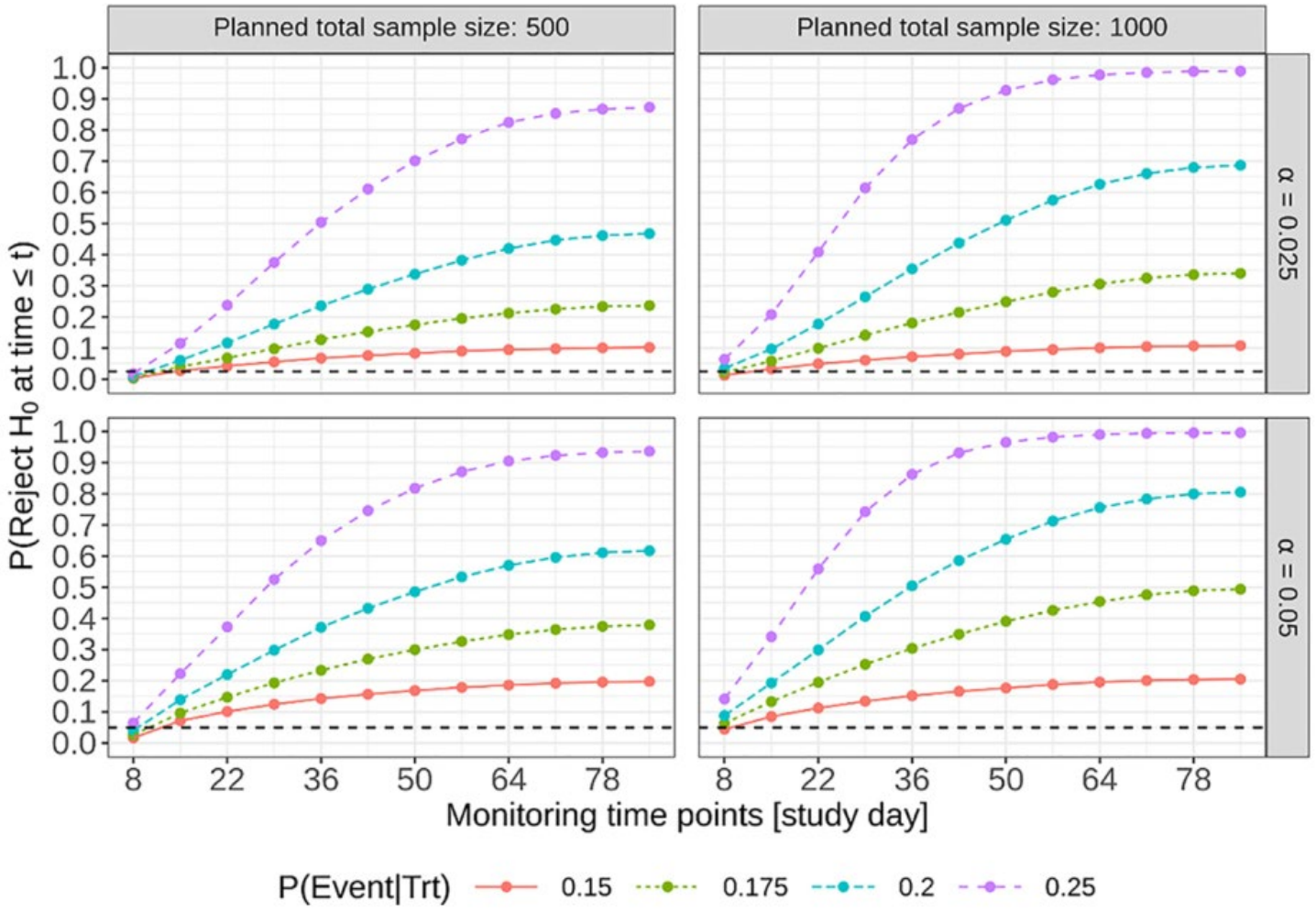
COVID-19 example: Simulation scenarios

Table 1
Specifications for the simulation study to assess the monitoring procedure's operating characteristics.

Parameter	Value
Uniform recruitment period	8 weeks
Sample size	$n = 500, 1000$
Treatment allocation	1:1
P(Event within 4 weeks| CTL)	0.15
P(Event within 4 weeks| TRT)	0.15, 0.175, 0.2, 0.25
Hazard ratio (as a result of assumptions above)	1, 1.18, 1.37, 1.77
One-sided significance level α	0.025, 0.05
Monitoring frequency	Weekly

- **Simulation scenarios** summarized in Table 1
- **Cox proportional hazards model** for analysis
- **Exponential distributions** for data generation
- **Monitoring procedure**: At each look a Wald test for hypotheses H0: HR=1 vs. H1: HR>1 at significance level α performed

COVID-19 example: Simulation results



- **Probability for rejecting the null hypothesis H0**
- P(Event|Trt)=0.15 is under H0, other scenarios under H1
- Note, corresponds to a group sequential design with **Pocock boundary** for which not the global type I error rate but the significance level at each time point is chosen. Therefore, **large sample approximation obtained through standard group sequential software**