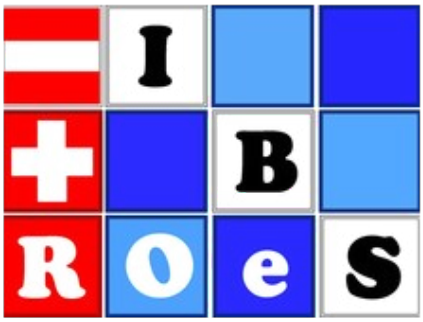
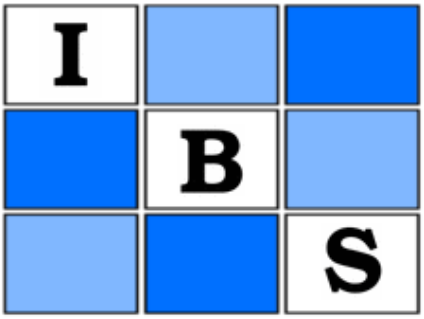
ADAPTIVE DESIGNS AND MULTIPLE TESTING PROCEDURES

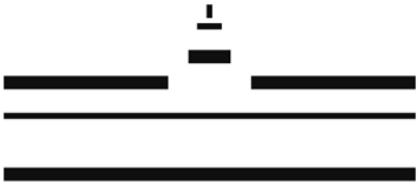
Workshop May 8 - 10, 2019

in Münster





Deutsche Region



WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER

Institut für Biometrie und Klinische Forschung

medizinische fakultät

Westfälische Wilhelms-Universität Münster

Organization committee:

René Schmidt (Münster), Robert Kwiecien (Münster),

Ellen Boldt (Münster), Andreas Faldum (Münster)

LOCATION

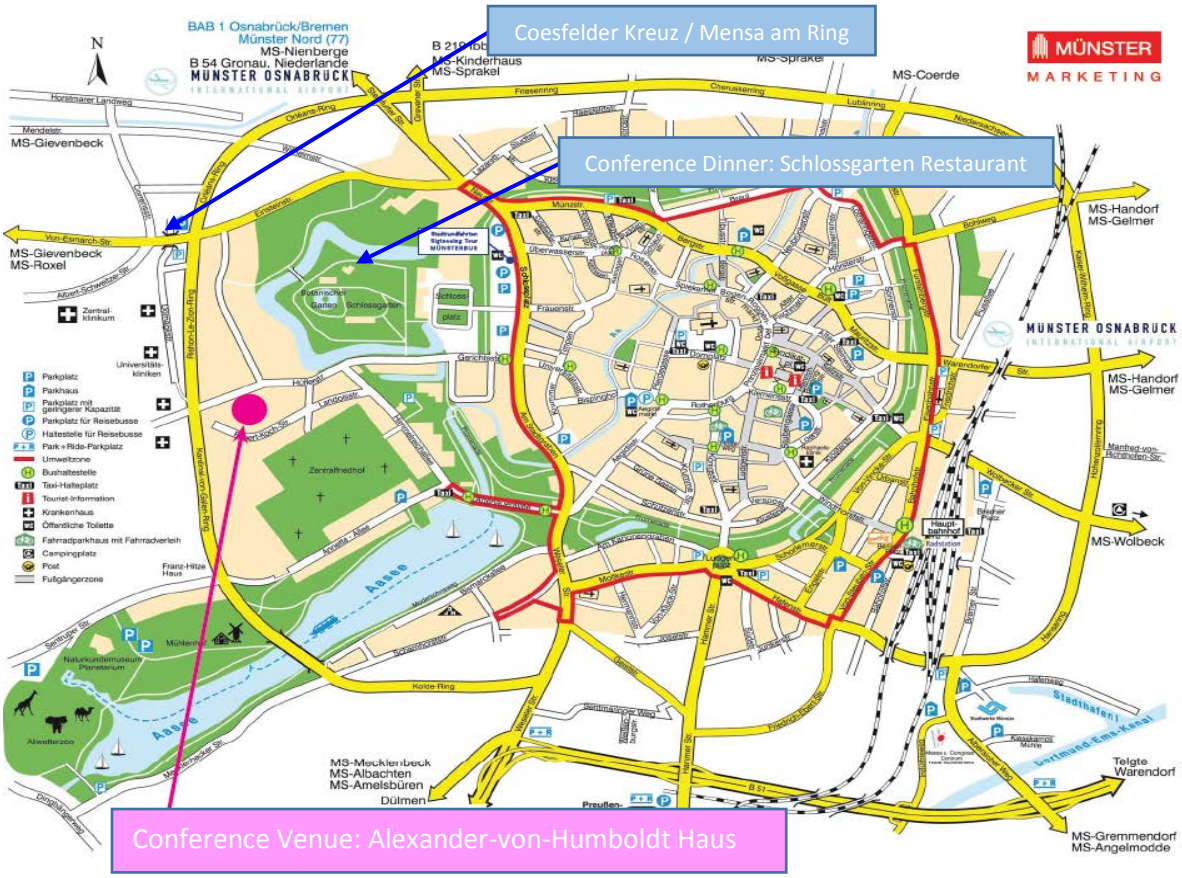
The conference takes place inside the Alexander-von-Humboldt Haus, located close to the Medical Faculty of the University of Münster:

Alexander-von-Humboldt Haus
Hüfferstrasse 61
48149 Münster

Arrival by car via highway A1 from direction north: Comming from the highway A1 from direction Bremen/Osnabrück, take exit "Münster-Nord" and drive in direction Münster (Steinfurter Straße/B 54). At the fourth traffic lights (Orleans-Ring) turn right, following direction "Unikliniken/Zoo". At the third traffic lights (Jungeblodtplatz) turn left into the Hüfferstrasse. After 150 m, the Alexander-von-Humboldt Haus is located on the right.

Arrival by car via highway A1 from direction south: Comming from the highway A1 from direction Köln/Dortmund, take exit "Münster-Süd" and drive in direction Münster (B 51). Follow the Weseler Strasse in direction Münster-Centrum. At the seventh traffic lights turn left onto the Kolde-Ring, following direction "Unikliniken/Zoo". At the fourth traffic lights (Jungeblodtplatz) turn right into the Hüfferstrasse. After 150 m, the Alexander-von-Humboldt Haus is located on the right.

Arrival by train and public transport: Take bus line 11, 14 or 22 from Münster Main Station (Hauptbahnhof) at bus stop C1. Leave the bus at Jungeblodtplatz. The Alexander-von-Humboldt Haus is located in short walking distance (150 meters back in direction Münster-Centrum).



CONFERENCE DINNER

There will be a conference dinner on

Thursday, May 9, 19:30

at the **Schlossgarten Restaurant** located in the backyards of the Palace of Münster in the neighbourhood of the Botanical Gardens (see City Map, page 1).

The address is:

Schloßgarten Restaurant
Schlossgarten 4
48149 Münster
Tel: 0251 - 987 96 96

https://www.schlossgarten.com/

LUNCH

Lunch can be taken at the **Mensa am Ring** located at the *Coesfelder Kreuz* (see City Map, page 1). It is located about 850 m from the workshop venue, reachable by foot in 10 minutes.

Mensa vouchers can be obtained at the conference office.

SCIENTIFIC PROGRAM OVERVIEW

Wednesday, May 08

11.00 Registration

12.00 – 13.30 Lunch

13.30 – 13.45 Welcome

13.45 – 15.15 Session *Research in Progress I* (Chair: Andreas Faldum)

- Carolin Herrmann, Maximilian Pilz, Meinhard Kieser, Geraldine Rauch: New and old ideas for optimizing sample size recalculation in two-stage adaptive group sequential study designs
- Cornelia Ursula Kunz, Nigel Stallard, Frank Fleischer: Parallel Adaptive Seamless Phase 2/3 Trials
- Tobias Mielke: Design considerations for sequential testing of nested subgroups

15.15 – 15.45 Coffee Break

15.45 – 17.15 Session *Research in Progress II* (Chair: Tobias Mielke)

- Robbie Peck: Optimal Decisions in the Portfolio Problem
- Amra Hot, Antonia Zapf: Adaptive Designs in Randomized Diagnostic Studies with Patient-Relevant Endpoints – First Considerations
- Thomas Burnett: Adding unplanned treatment arms to a Multi-Arm Multi-Stage study in progress

Thursday, May 09

9.00 – 10.30 Session *Adaptive Designs* I (Chair: Cornelia Ursula Kunz)

- Laura Kerschke, Andreas Faldum, Rene Schmidt: A New Log-Rank Test for Adaptive Survival Trials
- Franz König, Robin Ristl, Nicolás Ballarini, Heiko Götte, Armin Schüler, Martin Posch: Delayed treatment effects, treatment switching and heterogeneous patient populations: how to design and analyse randomized controlled trials in oncology
- Thomas Asendorf, Antonia Zapf, Christoph Anten, Tobias Mütze, Tim Friede: Blinded Sample Size Re-estimation for Negative Binomial Counts with Baseline Covariates

10.30 – 12.30 Coffee, discussions & small-group working on *Research in Progress*

12.30 – 14.00 Lunch

14.00 – 15.00 BMBF-Project ADIT: Selected Results (Chair: Robert Kwiecien)

- Robert Kwiecien, Andreas Faldum, Rene Schmidt: Constructing adaptive designs for one-sample-log-rank-test via stochastic integral processes
- Charlie Hillner, Werner Brannath, Kornelius Rohmeyer: Group sequential and adaptive designs with control of the „population-wise error rate“

15.00 – 15.30 Coffee Break

15.30 – 17.15 Honory Session (in German): Adaptive Designs und multiple Testprozeduren – Ein Rückblick und Ausblick anlässlich des 75. Geburtstags von Gerhard Hommel

- Andreas Faldum: „Gerhard Hommel: Erforsche das, was nicht interessant ist, nicht, nicht aber das nicht, was interessant ist!“
- Frank Bretz: How Gerhard Hommel taught me that power is not everything
- Peter Bauer: Das MULTAPTIV-Projekt, persönliche Erinnerungen und Ausblick

17.15 Meeting of the IBS-DR/ROeS Working Group on ADMTP

19.30 Conference Dinner at the Schloßgarten Café (see page 2)
Schlossgarten 4, https://www.schlossgarten.com/

Friday, May 10

09.30 – 11.00 Session *Multiple Testing* (Chair: Dennis Görlich)

- Arnold Janssen, Marc Ditzhaus: How to control the false discovery rate under dependence?
- Jonathan von Schröder, Thorsten Dickhaus: Efficient Calculation of the Joint Distribution of Order Statistics and Applications
- André Neumann, Thorsten Dickhaus: Non-parametric Archimedean generator estimation with implications for multiple testing

11.00 – 11.30 Coffee Break

11.30 – 12.30 Session *Adaptive Designs* II (Chair: Thomas Asendorf)

- Maximilian Pilz, Kevin Kunzmann, Carolin Herrmann, Geraldine Rauch, Meinhard Kieser: Adaptive two-stage designs with optimal performance characteristics
- Matthias Brückner, Andrew Titman, Thomas Jaki: Simulation-based Bayesian optimal adaptive designs for time-to-event models

12.30 End of workshop

ABSTRACTS (TALKS)

In alphabetic order by corresponding author

Blinded Sample Size Re-estimation for Negative Binomial Counts with Baseline Covariates

Thomas Asendorf[1], Antonia Zapf, Christoph Anten, Tobias Mütze, Tim Friede

[1] Institut für Medizinische Statistik, Universitätsmedizin Göttingen

Baseline covariates are frequently included in the primary analysis of clinical trials. Consequently, the planning of a clinical trial's sample size should also include effects of covariates on the primary endpoint. However, effects of covariates on primary endpoints are difficult to assess from previous studies, partially due to insufficient reporting, resulting in considerable uncertainty of the planned sample size. To reduce this uncertainty, we propose a blinded sample size re-estimation procedure for negative binomial distributed count data outcomes with baseline covariates, extending current methodology [1, 2]. Two methodologically different approaches for blinded estimation of nuisance parameters are introduced. The first approach is based on the expanded data set from Lyles et al. [3]. The second approach, derived from a likelihood based Wald-test, estimates the expected Fisher Information from blinded data. In a Monte Carlo simulation study, both methods are shown to appropriately adjust the sample size to maintain a pre-specified power, while not inflating the type I error rate. Finally, the proposed results are illustrated on an example and discussed.

References

[1] Friede T, Schmidli H. Blinded sample size reestimation with negative binomial counts in superiority and non-inferiority trials. Methods Inf Med. 2010;49(6):618–624.

[2] Friede T, Kieser M. Blinded sample size recalculation for clinical trials with normal data and baseline adjusted analysis. Pharm Stat. 2011;10(1):8–13.

[3] Lyles RH, Lin HM, Williamson JM. A practical approach to computing power for generalized linear models with nominal, count or ordinal responses. Stat Med. 2007;26(7):1632–1648.

Das MULTAPTIV-Projekt, persönliche Erinnerungen und Ausblick

Peter Bauer[1]

[1] Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna

Die Bereiche „Multiples Testen“ und „Adaptive Designs“ haben über die vergangenen Jahrzehnte besonderes Gewicht in der biometrischen Forschung in der Region Österreich-Schweiz und der Deutschen Region der Internationalen Biometrischen Gesellschaft erlangt. Anlässlich des 75. Geburtstags eines unserer Pioniere ist es naheliegend, sich an die Anfänge zu erinnern, aber risikoreich, angesichts der sich ändernden Landschaft biometrischer Forschung einen Ausblick zu wagen. Wenn man aber zur Altersklasse des Jubilars zählt, dann kann und braucht man später nicht mehr begründen, warum man mit dem Ausblick falsch lag.

How Gerhard Hommel taught me that power is not everything

Frank Bretz[1]

[1] Novartis Pharma AG

Gerhard Hommel and I have known each other for about 20 years, since the first regional MCP conference in Hannover. During this time I learned from him in many ways. In this presentation I look back on this time, share some aesthetic moments of our collaboration and explain why power is not everything. In particular, I will illustrate how working together on shortcut procedures has paved the way for graphical approaches to multiple testing, including some recent developments to visualize the truncated Hochberg procedure in serial and parallel gatekeeping settings using symmetric component graphs.

Simulation-based Bayesian optimal adaptive designs for time-to-event models

Matthias Brückner[1], Thomas Jaki

[1] Department of Mathematics and Statistics, Lancaster University

Design and analysis of late phase confirmatory clinical trials with time-to-event endpoints are dominated by frequentist methods. The need to specify a model for the hazard function, including prior distributions, and the lack of formal control of frequentist operating characteristics has limited the application of Bayesian approaches. We develop a general method for Bayesian decision-theoretic adaptive designs for time-to-event models which obey constraints on the type I and type II error even when arbitrary design adaptations are made by application of the conditional error principle. A simulation-based approach is used to obtain approximations of the expected utility and of frequentist error probabilities. We demonstrate the implementation of the general algorithm in a two-stage adaptive enrichment design with subgroup selection.

Adding unplanned treatment arms to a Multi-Arm Multi-Stage study in progress

Thomas Burnett[1]

[1] Mathematics and Statistics, Lancaster Unversity

Multi-Arm Multi-Stage (MAMS) trials are a powerful tool when choosing between multiple treatment options for further study. They allow for the comparison of multiple experimental treatments with a common control. Additionally, they allow the trial to be concluded early either for futility or efficacy if the evidence for a treatment or treatments becomes clear.

The generalised Dunnett method allows for the simultaneous hypotheses testing of the new treatments vs the control, while controlling the FamilyWise Error rate. This hypothesis testing structure allows for the early stopping of the trial, both for futility and efficacy, using pre-planned group sequential stopping boundaries. However, this does not allow for unplanned design modifications. Suppose a further new treatment becomes available during the progress of a MAMS study in progress, our goal is to incorporate this treatment into the trial while maintaining the FamilyWise Error Rate.

The conditional error approach allows unplanned design modifications while controlling the error rate at a pre-specified level for a given hypothesis test. We apply this principle to the Generalised Dunnett method by considering the conditional FamilyWise Error Rate under the global null. This allows updated group sequential stopping boundaries to be constructed that incorporate the additional treatment. Thus, this further new treatment may be added to the trial in progress without compromising its integrity.

New and old ideas for optimizing sample size recalculation in two-stage adaptive group sequential study designs

Carolin Herrmann[1], Maximilian Pilz, Meinhard Kieser, Geraldine Rauch

[1] Institute of Biometry and Clinical Epidemiology, Charité – Universitätsmedizin Berlin

A careful sample size calculation contributes substantially to the success of a clinical trial. However, the required parameter estimations do not always exist or are inconsistent across the literature. One way to tackle this problem is to adapt the sample size during an ongoing trial. Hence, the topic of sample size recalculation is one of the important aspects in adaptive group sequential study designs. Classical recalculation rules usually show only moderate performance properties with respect to (conditional) power and their average sample size. In this talk, new and old ideas on an optimized new sample size recalculation rule are presented. These ideas cover resampling arguments to address the underlying randomness of the interim result, continuity correction methods for a possible reduction of variability and external factors such as the local stopping boundaries, the futility boundary and the allowed maximal sample size.

Questions to the audience:

- How much does the definition of the futility stop bound add to a desirable sample size recalculation strategy?
- What is a reasonable maximal sample size that can be included in the planning assumptions for the simulations?

Group sequential and adaptive designs with control of the „population-wise error rate“

Charlie Hillner[1], Werner Brannath, Kornelius Rohmeyer

[1] Kompetenzzentrum für Klinische Studien Bremen (KKSB), University of Bremen

In confirmatory clinical trials that concern tests of several hypotheses in several populations the multiple type-I error is usually kept small by controlling the family-wise error rate (FWER). However, if a treatment or a treatment strategy is tested in several disjoint populations, each population is effected by only a single hypothesis test. In this case, the control of the FWER might be too conservative, so a more liberal multiple type I error rate, which we denote as “population-wise error rate (PWER)”, is considered. Suppose there are m possibly overlapping populations Pi in each of which the efficacy of a treatment strategy Ti is to be investigated by testing a hypothesis Hi. The population-wise error rate then describes the probability that a randomly selected future patient is assigned to an inefficient treatment strategy. We propose two-stage group sequential and adaptive design strategies that ensure control of this new error rate and will investigate their utility by means of examples. For group sequential designs, we will introduce two-stage test procedures for the special case of two overlapping populations, where the same treatment strategy is investigated in both populations. For adaptive designs, we propose an adoption of the CRP-principle by Müller and Schäfer to the PWER-approach.

Adaptive Designs in Randomized Diagnostic Studies with Patient-Relevant Endpoints – First Considerations

Amra Hot[1], Antonia Zapf

[1] Institute of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf

Diagnostic accuracy studies are performed to assess how well a diagnostic test can distinguish diseased and non-diseased individuals. However, the resulting diagnostic information is only beneficial if it is appropriately used in subsequent patient management decisions and thus clinically relevant outcomes, such as morbidity, mortality or health related quality of life, are improved in the long run. A joint evaluation of test and therapy regarding patient-relevant outcomes is the aim of randomized diagnostic trials.

Here, an important aspect is the sample size calculation, which is based on assumptions concerning the prevalence and the diagnostic accuracy, i.e. sensitivity and/or specificity, of the test as well as the suggested therapy effect from preliminary studies. Due to incorrect assumptions there is always a risk of incorrect estimation of the sample size and leading to an over- or underpowered trial.

Thus, adaptive trial designs for randomized diagnostic studies would be valuable, which allow modifications of the sample size or aspects of the study design by means of predetermined interim analyses. Until now, there is little research regarding adaptive designs for randomized diagnostic trials. As part of this work, the focus is to differentiate to what extent blinded interim evaluations concerning patient-relevant outcomes can be performed and when blinding is no longer fulfilled, without undermining the statistical validity and integrity of the trial. In the case of an unblinded or blinded interim analysis the type I error should be adjusted or not, particularly. The considering of the extent, to which statistical power and/or the type I error may be affected by a re-estimation of the sample size based on the prevalence, therapy effect or diagnostic accuracy of the test and whether a desired power is achieved at all, is a key question of this work.

In addition, another challenge is the variety of study designs proposed in the literature for randomized diagnostic studies that differ at time point of the randomization and their basic design. However, the nomenclature is not consistent. A clear distinction of these designs will contribute to the development of a sample size re-estimation formula for randomized diagnostic trials.

How to control the false discovery rate under dependence?

Arnold Janssen[1], Marc Ditzhaus

[1] Mathematical Institute, Heinrich-Heine-University Duesseldorf

It is a big issue to control the false discovery rate (FDR) for a collection of tests based on dependent test statistics. It is well known that traditional multiple procedures, for instance, the Benjamini and Hochberg test, may fail for negatively correlated statistics. On the other hand, the straight forward Bonferroni procedure always controls the FDR but it can be very conservative.
The present talk discusses multiple procedures for a general setting of dependent p-values. We propose new adaptive multiple tests with data dependent critical values which allow FDR control at a prespecified level. These multiple procedures are of interest for genome-wide association studies and they can be recommended in practice.

A New Log-Rank Test for Adaptive Survival Trials

Laura Kerschke[1], Andreas Faldum, Rene Schmidt

[1] Institute of Biostatistics and Clinical Research, University of Münster

The standard approach to compare the survival curves of two treatment groups is the two-sample log-rank test proposed by Mantel [1] and Peto and Peto [2]. It has favorable properties, as it is optimal under proportional hazards and can easily be generalized to handle more than two groups. However, generalization of the log-rank test to multiple groups a priori only allows for a global comparison of study arms and multiplicity adjustment is required if several pairwise comparisons are made. Using a Bonferroni correction results in a power loss, since dependencies among test statistics are ignored. Strategies that account for the correlation structure of the two-sample log-rank test statistics have, so far, only been proposed for single-stage, fixed sample settings. For multi-arm multi-stage designs these strategies cannot be transferred readily. This is due to the fact that the common two-sample log-rank test statistic is obtained from pooled data of both treatment groups and cannot be written as the difference of two independent random variables derived from non-overlapping populations.
To overcome this issue, we propose an alternative two-sample log-rank test such that the underlying test statistic is similar to that of an unpaired z-test with known variance. On this basis the well-known methodology for comparisons of means can immediately be transferred to the survival setting. The proposed method is compared to the common two-sample log-rank test by simulation based on a two-step drop-the-loser design.

References

[1] Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports. 50, 163-170.

[2] Peto, R. and Peto J. (1972). Asymptotically efficient rank invariant test procedures. Journal of the Royal Statistical Society, Series A. 135, 185-207.

Delayed treatment effects, treatment switching and heterogeneous patient populations: how to design and analyse randomized controlled trials in oncology

Franz König[1], Robin Ristl, Nicolás Ballarini, Heiko Götte, Armin Schüler, Martin Posch

[1] Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna

Comparing survival distributions based on censored data is in general challenging since conclusions are sought about distributions which are observed incompletely. In medical statistics, and in particular in oncology, the logrank test and the Cox model have been established as key tools, which do not require specific distributional assumptions. Under the relatively weak assumption of proportional hazards, they are efficient and their results can be interpreted unambiguously. However, delayed treatment effects, disease progression, treatment switchers or the presence of subgroups with differential treatment effects may challenge the assumption of proportional hazards. In practice, weighted logrank tests emphasizing either early, intermediate or late event times via an appropriate weighting function may be used to accommodate for an expected pattern of non-proportionality.

We model these sources of non-proportional hazards via a mixture of survival functions with piecewise constant hazard. The model is then applied to study the power of unweighted and weighted log-rank tests, as well as maximum tests allowing different time dependent weights. Simulation results suggest a robust performance of maximum tests across different scenarios, with little loss in power compared to the most powerful among the considered weighting schemes and huge power gain compared to unfavorable weights.

We further propose a framework to perform an interim analysis and calculate the conditional power under the non-proportional hazards model to allow futility stopping decision making, sample size reassessment or modification of the testing procedure.

Parallel Adaptive Seamless Phase 2/3 Trials

Cornelia Ursula Kunz[1], Nigel Stellard, Frank Fleischer

[1] Global Biostatistics and Data Sciences, Boehringer Ingelheim

As the number of new compounds approved by regulatory agencies remains low, new statistical methods are needed in order to accelerate the drug development process. Recent statistical methodology has focused on combining different phases to shorten the time to approval of a new drug. In particular, seamless adaptive Phase 2/3 designs have gained popularity among pharmaceutical companies. If conducted in an inferentially seamless way, these designs allow reductions in sample size for a Phase 3 pivotal trial as they combine the data obtained in Phase 2 with data obtained in the Phase 3.

However, before a new drug can be approved by, for example, the FDA, the manufacturer has to provide substantial evidence of the effectiveness of the new treatment. Usually, this is interpreted as conducting at least two pivotal trials which both have to be significant [1].

As the second pivotal trial can only be started after the Phase 2 part of the seamless trial has finished, the overall development time is often unaffected.

Shun et al. [2] have undertaken preliminary work on the difference between conducting one versus two trials in Phase III. However, they only focus on single stage trials. We therefore extend their considerations to more flexible trial approaches such as adaptive seamless Phase 2/3 designs where the aim is to combine one Phase 2 with one Phase 3 trial [3]. We investigate several different ways to combine the data obtained in Phase 2 with the data from the two Phase 3 trials.

By deriving expressions for frequentist error rates and sample sizes for innovative design approaches, optimal design strategies can be considered, both for a single homogeneous population and when there are differences between the patient populations in which different trials are conducted.

References

[1] FDA (1998): Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products.

[2] Shun Z, Chi E, Durrleman S and Fisher L (2005): Statistical consideration of the strategy for demonstrating clinical evidence of effectiveness—one larger vs two smaller pivotal studies. Statistics in Medicine, 24:1619-1637.

[3] Kunz CU, Friede T, Parsons N, Todd S and Stallard N (2015): A comparison of methods for treatment selection in seamless phase II/III clinical trials incorporating information on short-term endpoints. Journal of Biopharmaceutical Statistics, 25:170-198.

Constructing adaptive designs for one-sample-log-rank-test via stochastic integral processes

Robert Kwiecien[1], Andreas Faldum, René Schmidt

[1] Institute of Biostatistics and Clinical Research, University of Münster

In financial mathematics, stochastic integral processes usually represent the value of assets and portfolios in the course of time. The stochastic integrand represents the actions (="adaptations") of the corresponding shareholder. In an adaptive Design, we change the structure of the test statistic and analysis method in a data driven way in the course of an experiment or a trial.

We will demonstrate that an approach from the financial mathematics also works for some test statistic processes regarding adaptive designs, where the integrand represents the adaptation rules of these designs. In terms of the one-sample Log-Rank test, the adaptive actions during a corresponding trial, are still limited, as well as the mathematical assumptions are limiting the range of applications. The structure of an Ito-integral process, which involves a predictable process in terms of an integrand, can be used to implement adaptation rules directly in the construction of a test statistic process, which are essentially only limited to 'use information from the past'.

We will show how to characterize the (asymptotic) distribution of those test statistic processes in a comfortable way, and how to construct adaptive test sequences, which controls the level of significance, and allows very flexible adaptations.

We will present a part of the research results of the project 'Adaptive Designs in individualized therapy - ADIT', FKZ 01EK1503A, which is sponsored by the Federal Ministry of Education and Research.

Design considerations for sequential testing of nested subgroups

Tobias Mielke[1]

[1] QS Consulting, Janssen Pharmaceuticals

Clinical studies including multiple subgroups require appropriate testing strategies for the control of the family wise error rate. Standard Bonferroni adjustments will assert control of the family wise error rate but are conservative for positively correlated test statistics. A variety of approaches exist to utilize the known correlation structure in testing effects on multiple study populations (see Ondra et al.(2016), section 3.1). Spiessens & Debois (2010) generalized group sequential testing ideas to applications of subgroup analyses in clinical trials. The full significance level α is split to the different subgroups under consideration. Adjusted critical values for the Spiessens & Debois approach are directly transferred from standard group-sequential trials to the problem of testing subgroups. While the proposed subgroup analyses follow the same mechanisms of group-sequential trials, the development of study designs using the approach of Spiessens & Debois will differ in an important aspect from standard group-sequential study designs. Subgroups under consideration will typically have different underlying treatment effects. Optimized α-spending strategies can take assumptions on the treatment effects and the correlation of the sub-populations into account. An additional optimization problem may exist when subgroups are defined based on continuous biomarkers. Different biomarker cutoffs for the definition of subgroups will lead to different prevalences and mean effects of the subgroups under consideration. An optimized biomarker cutoff together with an optimized α-spending strategy could increase the power for testing effects in subgroups. However, these optimization targets will typically depend on underlying effect assumptions. Deviations from these effect assumptions may decrease power of the study. Combination test ideas (e.g. Zhao et al. (2010)) combined with the motivation of the MCPMod approach (Bretz et al. (2005), Thomas et al. (2019)) allow to define optimized subgroup testing approaches, which may account for the uncertainty on the subgroup effects (and even more generally for the definition of subgroups). These approaches will require adequate multiplicity corrections for the control of the family wise error rate. In this presentation, different approaches for the selection of biomarker cutoffs and for the sequential testing of effects in nested subgroups will be presented. Limitations and benefits of the presented approaches will be discussed on a case study.

References

[1] Ondra, T., Dmitrienko, A., Friede, T., Graf, A., Miller, F., Stallard, N. and Posch, M. Methods for identification and confirmation of targeted subgroups in clinical trials: A systematic review. Journal of Biopharmaceutical Statistics, 26, pp.99-119 (2016)

[2] Spiessens, B., and Debois, M. Adjusted significance levels for subgroup analyses in clinical trials. Contemporary Clinical Trials, 31, pp. 647-656 (2010)

[3] Zhao, Y.D., Dmitrienko, A. and Tamura, R. Design and Analysis Considerations in Clinical Trials With a Sensitive Subpopulation. Statistics in Biopharmaceutical Research, 2, pp. 72-83 (2010)

[4] Bretz, F., Pinheiro, J.C. and Branson, M. Combining Multiple Comparisons and Modeling Techniques in Dose Response Studies. Biometrics, 61, pp. 738-748 (2005)

[5] Thomas, M., Bornkamp, B., Posch, M. and König, F., A multiple comparison procedure for dose-finding trials with sub-populations. https://arxiv.org/abs/1811.09824v1 [stat.ME] (2018)

Non-parametric Archimedean generator estimation with implications for multiple testing

André Neumann[1], Thorsten Dickhaus

[1] Institute for Statistics, University of Bremen

In multiple testing, the family-wise error rate can be bounded in some situations by the copula of the test statistics. Assuming that this copula is Archimedean, we consider two non-parametric Archimedean generator estimators. More specifically, we use the non-parametric estimator from Genest, Neslehovà and Ziegel (2011) and a slight modification thereof. In simulations, we compare the resulting multiple tests with the Bonferroni test and the multiple test derived from the true generator as baselines.

Optimal Decisions in the Portfolio Problem

Robbie Pack[1]

[1] Mathematical Sciences, University of Bath

Suppose a trial sponsor has a portfolio of various drugs in phase II trials which if successful will become available for phase III trials in the near future. We consider the portfolio problem, which deals the optimal allocation of a R&D budget to phase III confirmatory trials within some planning horizon. We require a decision strategy to allocate the overall budget to the available drugs for phase III trials in order to maximise the expected net present value of the portfolio. This decision strategy must specify the optimal phase III sample sizes, given the remaining budget and drug parameters.

Previous approaches to this problem have used integer programming formulations or simulation models. We use a dynamic programming approach to derive the optimal decisions, which may scale up more efficiently in computational workload as the portfolio considered becomes more complex.

The use of group sequential designs which allow early stopping can benefit the portfolio by firstly increasing the time one can market the drug until patent expiry, and secondly allowing the reinvestment of saved resources back into the portfolio. We analyse the benefit to the portfolio value and changes to the optimal decision rules when using group sequential methods.

Adaptive two-stage designs with optimal performance characteristics

Maximilian Pilz[1], Carolin Herrmann, Geraldine Rauch, Meinhard Kieser

[1] Institute of Medical Biometry and Informatics, University of Heidelberg

In clinical trials, the choice of an adequate sample size is a crucial issue. While traditionally a design with a fixed sample size is applied, flexible strategies with one or several interim analyses are becoming increasingly popular.
In an adaptive two-stage design, one has to specify the first- and the second-stage sample sizes as well as boundaries for early stopping for futility and efficacy and the critical value for the final analysis. There exists a variety of approaches on how to choose these key figures. However, from the viewpoint of design efficiency, one can regard all these design elements as tuning parameters for a given optimization problem. Solving this problem yields to an optimal adaptive design. For its construction, neither a specific sample size recalculation rule nor a combination test has to be presumed. To guarantee type one error rate control and a sufficiently large power, these constraints are included as inequalities in the optimization framework.
Up to now, our research focused on optimizing expected sample size under a point alternative hypothesis. In our recent work, the impact of the chosen optimality criterion is investigated. For instance, the expected sample size under the alternative hypothesis or under the null hypothesis as well as a mixed criterion consisting of expected sample size and conditional power are considered as quantities to be optimized.
Furthermore, Bayesian priors for the effect size will be incorporated and the resulting optimal designs for the use of predictive power instead of classical conditional power will be analyzed.
The results give a deep insight in the behavior of optimal adaptive designs and their dependence on the applied optimality criterion and the prior distribution.
The methodology will be illustrated by clinical trial examples covering a variety of effect sizes. The proposed method is highly relevant for clinical trial practice because the derived rules could directly be used for planning a clinical trial and, due to their optimality, they are superior to all other possible design choices.

Efficient Calculation of the Joint Distribution of Order Statistics and Applications

Jonathan von Schroeder[1], Thorsten Dickhaus

[1] Institute for Statistics (FB 3), University of Bremen

We consider the problem of computing the joint distribution of order statistics of stochastically independent random variables in one- and two-group models. While recursive formulas for evaluating the joint cumulative distribution function of such order statistics exist in the literature for a longer time, their numerical implementation remains a challenging task. We tackle this task by presenting novel generalizations of known recursions which we utilize to obtain exact results (calculated in rational arithmetic) as well as faithfully rounded results. Finally, some applications in stepwise multiple hypothesis testing are discussed.