

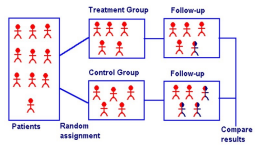
## *Does Randomization matter in RD clinical trials?*

Ralf-Dieter Hilgers  
rhilgers@ukaachen.de  
Institut für Medizinische Statistik  
RWTH-Aachen Universität



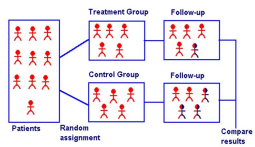
This research is part of the EU-FP7 IDeA project (GA No. 602552) and embedded in DH's EU Horizon 2020 fundings from EJP RD (GA No. 825575) and ERICA (GA No 964908)



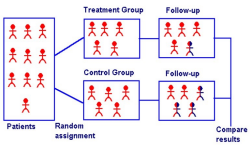


# Content

- Background
- Randomization: What it is – What it is not!
- What randomization can help us to do or should help us?
- What are elements of a structured choice approach?
- Which designs are developed right now?
- Final remarks



# BACKGROUND



# Background Knowledge – Learning Objectives

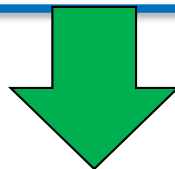


I have a rough idea

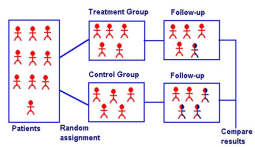
- how treatment allocation works in clinical trials
- what random allocation is

I will understand

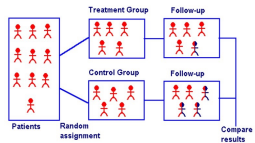
- ... the value of different randomization procedures
- ... that no randomization procedures fits all purposes
- ... how to select a randomization procedure based on scientific arguments
- ... the importance in RD trials



*We are not dealing with other methods to reduce the impact of bias!*



# **RANDOMIZATION: WHAT IT IS – WHAT IT IS NOT!**



# Randomization: *Historical examples*

Van Helmont, 1662

## Lotteries in health care

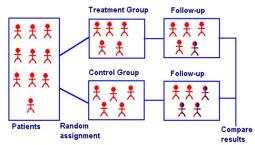
Lottery has been used and is still used to ensure fairness in health care. In the 17th century, to settle a dispute he was having with orthodox practitioners who used bloodletting and purging for treatment, the Flemish physician John Baptiste Van Helmont made the following proposition: “Let us take out of the hospitals ... 200 or 500 poor people, that have fevers, pleurisies. Let us divide them into halves, let us cast lots, that one halfe of them may fall to my share, and the other to yours; I will cure them without bloodletting and sensible evacuation; but you do, as ye know.... We shall see how many funerals both of us shall have.”<sup>6</sup>

MRC, 1948

The randomised trial of streptomycin also illustrates how lottery has been used to distribute limited supplies of a potentially beneficial intervention.<sup>10</sup> This

### Summary points

- Casting or drawing of lots has been used for thousands of years to help deal with uncertainty and ensure fairness
- It was proposed in the 17th century and adopted in the 19th century for making fair comparisons between alternative medical treatments
- It has also been used for the fair distribution of limited resources
- It is a fair way of distributing the hoped for benefits and unknown risks of inadequately evaluated forms of health care



# Randomization in Guidelines: What it is!

## 2.3.2 Randomisation

Randomisation introduces a deliberate element of chance into the assignment of treatments to subjects in a clinical trial. During subsequent analysis of the trial data, it provides a sound statistical basis for the quantitative evaluation of the evidence relating to treatment effects. It also tends to produce treatment groups in which the distributions of prognostic factors, known and unknown, are similar. In combination with blinding, randomisation helps to avoid possible bias in the selection and allocation of subjects arising from the predictability of treatment assignments.

ICH E9, p12

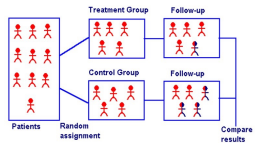
### a) Randomisation

In conducting a controlled trial, randomised allocation is the preferred means of assuring comparability of test groups and minimising the possibility of selection bias.

ICH E8



**Random Allocation -> stat. Test, balance cov., avoid bias**



# Randomization: What is is not!

The NEW ENGLAND JOURNAL of MEDICINE

*n engl j med 2020 382*

SOUNDING BOARD

## The Magic of Randomization versus the Myth of Real-World Evidence

In generalizing the results of a randomized trial, the assumption is not that the patient population studied is representative of all patients but rather that the proportional effects of the treatment studied on each specific health outcome should be similar in different circumstances, unless there is good reason to expect otherwise.<sup>15</sup> Consequently, valid estimates of the absolute benefits and harms of a treatment can be obtained by applying reliable randomized evidence for its separate proportional effects on each outcome of interest to the absolute incidence of these outcomes in observational studies conducted within a particular population. For

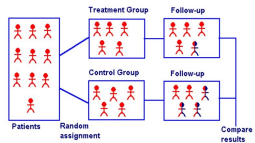
), F.R.C.P., Martin Landray, Ph.D., F.R.C.P.,  
d Peto, F.R.S.

However, because of the potential biases inherent in observational studies, such studies cannot generally be trusted when — as is often the case — the effects of the treatment of interest are actually null or only moderate (i.e., less than a twofold difference in the incidence of the health outcome between using and not using the treatment).<sup>4-6</sup> In those circumstances, large obser-



**Random Allocation is not Random Sampling**





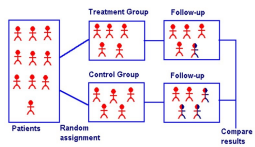
# Random Allocation Procedures (Examples and Properties)

- CR** *Complete randomization*: probability that patient  $i$  will receive treatment E is always 0.5
- RAR** *Random Allocation Rule*: randomize so that half of the  $n$  patients receive treatment E
- PBR( $m_s$ )** *Permuted Block Randomization*: Implementation of RAR within  $k$  Blocks of size  $m_s$  ;  $1 \leq s \leq k$
- BSD( $b$ )** *Big Stick design*: CR allow for imbalance within a limit  $b$
- MP( $b$ )** *Maximal Procedure*: Impose uniform probability to all sequences allowing for imbalance within a limit  $b$ ,
- EBC( $p$ )** *Efron's Biased Coin*: flip a biased coin ( $p$ ) in favour of the treatment which is allocated less frequently
- Chen( $p, b$ )** *Chen's design*: flip a biased coin ( $p$ ) in favour of the less frequently allocated treatment allowing for imbalance within a limit  $b$

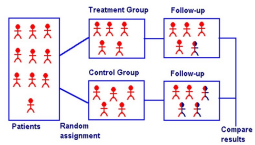
...



**Various procedures can be used for Random Allocation**



# WHAT RANDOMIZATION CAN HELP US TO DO OR SHOULD HELP US?



# Awareness of Randomization

## What the theory tells us:

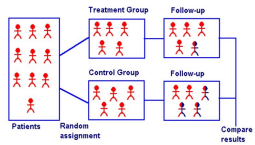
- no randomization procedure performs best with all criteria, Rosenberger (2016), Atkinson (2014),...

## What (applied) scientist mostly feel about randomization is ....

- scepticism, is a „*must*“
- that the *principle is unclear*
- that it is just *allocation* and *unequal group size* is a major problem
- that it is for *balancing covariates* but does mostly not work
- that selection of a procedure is by opinion or software availability

## What the literature mirrors is ...

- there is less or no training in randomization (necessary)
- there is no recommendation to give scientific arguments for the choice of randomization procedure, neither ICH Guidelines nor CONSORT Statement



# What Randomization should help us?

Berger et al. *BMC Med Res Methodol* (2021) 21:168  
<https://doi.org/10.1186/s12874-021-01303-z>

BMC Medical Research  
Methodology

RESEARCH

Open Access

## A roadmap to using randomization in clinical trials



Vance W. Berger<sup>1</sup>, Louis Joseph Bour<sup>2</sup>, Kerstine Carter<sup>3</sup>, Jonathan J. Chipman<sup>4,5</sup>, Colin C. Everett<sup>6</sup>, Nicole Heussen<sup>7,8</sup>, Catherine Hewitt<sup>9</sup>, Ralf-Dieter Hilgers<sup>7</sup>, Yuqun Abigail Luo<sup>10</sup>, Jone Renteria<sup>11,12</sup>, Yevgen Ryznyk<sup>13</sup>, Oleksandr Sverdlov<sup>14</sup> and Diane Uschner<sup>15</sup> for the Randomization Innovative Design Scientific Working Group

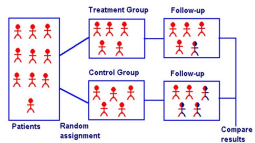
- .... mitigate selection bias due to an investigator's potential to selectively enroll patients into the study
- .... tendency to promote similarity of treatment groups with respect to known and unknown confounders
- ....an important role in statistical analysis of the clinical trial.

**Table 1** Considerations for the choice of a restricted randomization procedure

Objective	Desired feature(s) of a randomization procedure
Mitigate potential for selection bias	A procedure should have high degree of randomness.
Mitigate potential for chronological bias.	A procedure should balance treatment assignments over time.
Valid and efficient treatment comparison	A procedure should have established statistical properties, provide strong control of false positive rate and yield unbiased, low variance estimates of the treatment difference. A procedure should preserve the unconditional allocation ratio (e.g. 1:1) at every allocation step and achieve approximately or exactly the target sample sizes per group.
Ease of implementation	Validated statistical software for implementing a randomization procedure must be in place.



**Random Allocation can mitigate selection bias**



# Let us talk about bias

## Bias....

### Allocation

if investigators know or predict which intervention the next eligible participant is supposed to receive (*syn. Selection Bias*)

→ may influence the way investigators approach potentially eligible participants and how they are assigned to the different groups, thereby selecting participants with good prognoses (i.e. anticipated good outcomes and treatment responses) into one group more than another.

*(Preventive: allocation concealment, e.g. the block sizes for randomization should not be known)*

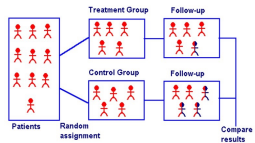
### Chronological

study participants allocated earlier to an intervention are subject to different response from participants who are recruited later.

*(Preventive: using small block sizes can reduce chronological bias, but must be balanced against the resulting risk of selection bias. )*

**+ Validity of the statistical inference model**

Catalogue of Bias Collaboration. Spencer EA, Heneghan C, Nunan D. **Allocation bias**. In: Catalogue of Bias 2017.



# Does Randomization the job?

## Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial

Craig M McDonald, Craig Campbell, Ricardo Erazo Torricelli, Richard S Finkel, Kevin M Flanigan, Nathalie Goemans, Peter Heydemann, Anna Kaminska, Janbernd Kirschner, Francesco Muntoni, Andrés Nascimento Osorio, Ulrike Schara, Thomas Sejersen, Perry B Shieh, H Lee Sweeney, Haluk Topaloglu, Már Tulinius, Juan J Vilchez, Thomas Voit, Brenda Wong, Gary Elfring, Hans Kroger, Xiaohui Luo, Joseph McIntosh, Tuyen Ong, Peter Riebling, Marcio Souza, Robert J Spiegel, Stuart W Peltz, Eugenio Mercuri, the Clinical Evaluator Training Group\*, and the ACT DMD Study Group\*

*Lancet, 2017*

### Summary

**Background** Duchenne muscular dystrophy (DMD) is a severe, progressive, and rare neuromuscular, X-linked recessive disease. Dystrophin deficiency is the underlying cause of disease; therefore, mutation-specific therapies aimed at restoring dystrophin function are being developed. This study evaluated the efficacy and safety of ataluren in patients with nonsense mutation Duchenne muscular dystrophy.

**Method** This multicentre, randomised, double-blind, placebo-controlled, phase 3 trial was conducted in 18 countries. Patients with nonsense mutation Duchenne muscular dystrophy were randomly assigned (1:1), via permuted block randomisation (block size of four) using an interactive voice-response or web-response system, to receive placebo or ataluren. Randomisation was stratified by age (<9 years vs ≥9 years), duration of previous corticosteroid use (6 months to <12 months vs ≥12 months), and baseline 6MWD (<350 m vs ≥350 m). Patients and caregivers were masked to group allocation until after database lock. The primary endpoint was change in 6MWD from baseline to week 48. We additionally did a prespecified subgroup analysis reflective of anticipated rates of disease progression. The study is registered with ClinicalTrials.gov, number NCT01471511.

**Results** 220 patients were randomly assigned (1:1) to placebo (n=110) or ataluren (n=110). The primary endpoint was change in 6MWD from baseline to week 48. We found no significant difference in 6MWD between the placebo and ataluren groups. The prespecified subgroup analysis showed no significant difference in 6MWD between the placebo and ataluren groups. The study was stopped because of slow recruitment.

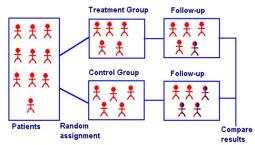
in 6MWD with at least 85% power ( $\alpha=0.05$ ). With the assumption that roughly 5% of patients would discontinue prematurely, a total of 220 patients (n=110 per group) would need to be enrolled.

Stratified PBR(4)

220 patients

55 Blocks

55 Blocks not divisible by  $2^3=8$  combinations (Stratification)



# Let's consider PBR(4)

PBR(4) →

1. **AAPP**
2. **APAP**
3. **APPA**
4. **PPAA**
5. **PAPA**
6. **PAAP**

Patients												
Strata	1	2	3	4	5	6	7	8	9	10	11	12
Z1	A	A	P	P	P	A	A	P	P	A	P	A
Z2	A	P	P	A	A	P	A	P	A	A	P	P
Z3	A	P	A	B	B	B	A	A	A	A	P	P
Z4	P	A	B	A	A	A	P	P	P	A	A	P
Z5	A	A	P	P	P	A	A	P	P	A	P	A
Z6	A	P	P	A	A	P	A	P	A	A	P	P
Z7	A	P	A	B	B	B	A	A	A	A	P	P
Z8	P	A	B	A	A	A	P	P	P	A	A	P

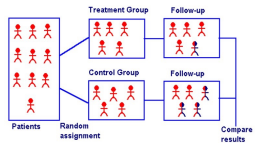
Adding *Stratification* to the randomization process **increases predictability** while reducing randomization list and thus **may increase potential for bias**

Knowledge about PBR(4)

- **60** blocks of size 4 to reach **240** patients
- **Between 60 and 120** allocations predictable = deterministic

*It is possible to vary the block length, again at random, perhaps using a mixture of blocks of size 2, 4, or 6.*

D G Altman 1, J M Bland  
How to randomise  
BMJ 1999 :703-4



# What is about PBR(6)?

**PBR(6) →**

**1. AAAPPP**

**2. APAAPP**

**3. ...**

*(15 Seq.)*

Range or **deterministic allocations** =  
 [number of blocks; half blocksize times number of blocks]

Knowledge about PBR(6)

→ **40** blocks of size 6 to reach **240** patients

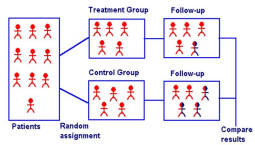
→ **Between 40 and 120 allocations predictable = deterministic allocation**



**Larger Blocksizes offer a reduced number of deterministic allocations.  
 Deterministic allocations is a simple (indirect) metric for allocation / selection bias.**

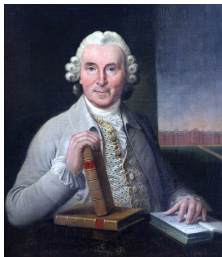
*Random Allocation Rule improves upon PBR's with respect to deterministic allocations.*





# Is it really important to....

- a) reach the planned number of allocations to treatment groups ?
- b) balance heterogeneity in patients by stratification ?



## Treatment comparisons must be fair

*Untrustworthy treatment comparisons are those in which biases, or the play of chance, or both result in misleading estimates of the effects of treatments. Fair treatment comparisons avoid biases and reduce the effects of the play of chance.*

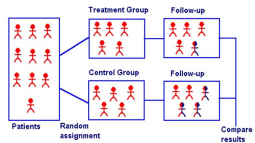
<https://www.jameslindlibrary.org/research-topics/fair-tests-of-treatments/treatment-comparisons-must-be-fair/>



**Predictability → (Allocation) Bias**

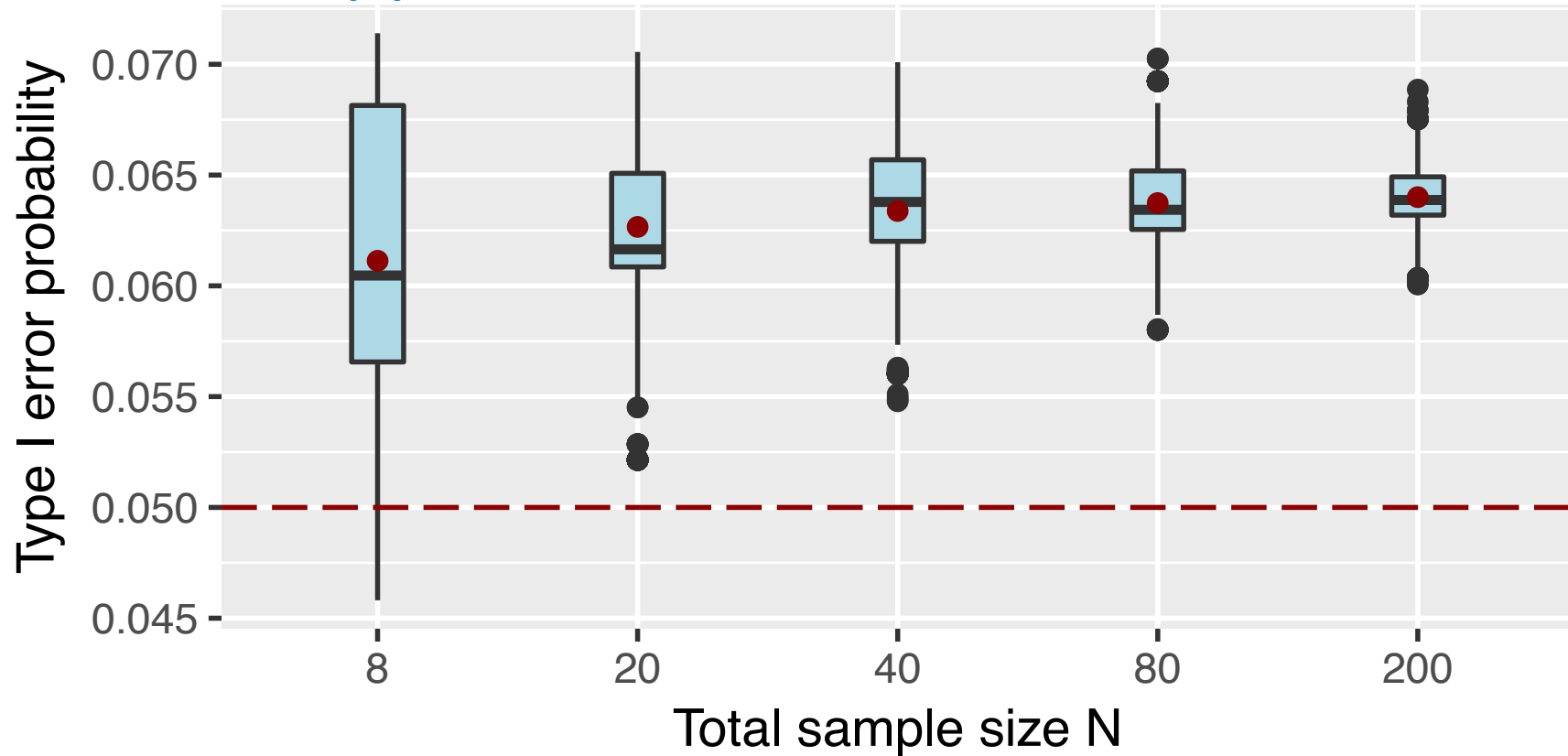


**Credibility → Validity of Trial Results → Reproducibility**



# Quantification of Allocation Bias versus N

PBR(4); allocation bias effect = 25% of effect size



Number of allocation sequences

36

7776

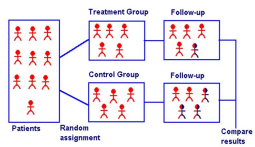
$6 \cdot 10^7$

$3.6 \cdot 10^{15}$

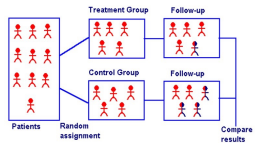
$8 \cdot 10^{38}$

ICH E9, p12

approaches. The interpretation of statistical measures of uncertainty of the treatment effect and treatment comparisons should involve consideration of the potential contribution of bias to the p-value, confidence interval, or inference.



# WHAT ARE ELEMENTS OF A STRUCTURED SELECTION APPROACH FOR A RANDOMIZATION PROCEDURE?



# ERDO template

## Evaluation of Randomization Procedures for Trial Design Optimization



ERDO

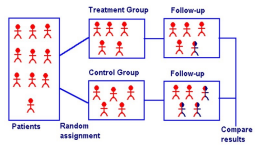
- Introduction
- Objective
- **Framework**
- Evaluation Methods
- Results and Decision
- Discussion and Clinical Implication
- Conclusion



- Assumptions – design, clinical settings
- Options – suitable set of RPs
- Metrics – evaluation criterion



**Benchmark process of the choice of the RP**



## ERDO - Objective

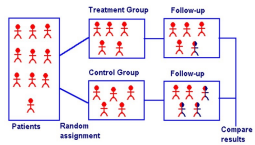
What we learned so far:

*It would be surprising, if a randomization procedure would be „optimal“ for all settings:*

**To quantify the effect of allocation bias** on the trial result (e.g. on the p-value) **the clinical situation** (primary endpoint variable, minimal effect one would not like to overlook → sample size; design specification of the trial) **has to be taken into account.**



**The problem of selecting the appropriate randomization procedure should be described, taking into account the particular situation specific to the clinical trial.**



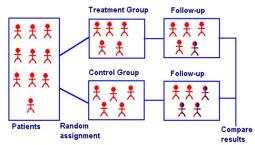
# ERDO - Clinical Design Setting

## Specify

- primary **endpoint** variable
- minimal effect one would not like to overlook → **sample size**
- (number of) **treatment arms**
- **design** (parallel group, crossover, etc. // stratification)
- **Adaptation**
- ...



**Specify the design characteristics**



# ERDO – Specify the Bias Model

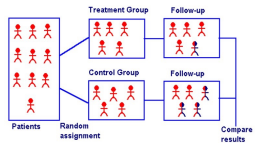
Effect on	Measure	$i^{\text{th}}$ patient ; $D_{i-1} = N_E(i-1) - N_C(i-1)$
Predictability (direct)	<ul style="list-style-type: none"> <li>Allocation bias</li> </ul>	<ul style="list-style-type: none"> <li><math>\tau_i(\eta) = \eta \operatorname{sgn}(D_{i-1})</math></li> <li><math>\tau_i(\eta_j)</math>, <math>j^{\text{th}}</math> center number</li> <li><math>\tau_i(\eta_j)</math>, <math>j^{\text{th}}</math> stage number</li> <li><math>\tau_i(\eta_j)</math>, <math>j^{\text{th}}</math> endpoint number</li> <li><math>\tau_i(\eta(t))</math>, <math>\eta(t)</math> time dependent</li> </ul>
Confounding (time trend)	<ul style="list-style-type: none"> <li>Chronological bias</li> </ul>	<ul style="list-style-type: none"> <li><math>\tau_i(\vartheta) = i \vartheta</math></li> <li><math>\tau_i(\vartheta) = \log\left(\frac{i}{N}\right) \vartheta</math></li> <li><math>\tau_i(\vartheta) = \mathbf{I}_{\{i \geq n_0\}} \vartheta, n_0 \leq N</math></li> </ul>



**Bias model specifications**

Extension:

- Additive allocation and chronological bias
- Combined unified assessment criterion (normalization with Derringer-Suich function)



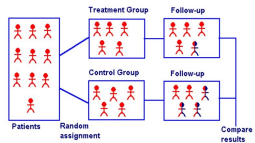
# ERDO – Specify the Evaluation Criterion

Effect on	Measure	$i^{\text{th}}$ patient ; $D_i = N_E(i) - N_C(i)$
Power	<ul style="list-style-type: none"> <li>Power due to imbalance</li> <li>Imbalance/loss</li> </ul>	<ul style="list-style-type: none"> <li><math>Power(D_N)</math></li> <li><math>D_i \frac{(D_N)^2}{N}</math></li> </ul>
Predictability (indirect)	<ul style="list-style-type: none"> <li>Correct Guesses</li> </ul>	<ul style="list-style-type: none"> <li><math>E(G)</math>, where G is number of correct guesses</li> </ul>



approaches. The interpretation of statistical measures of uncertainty of the treatment effect and treatment comparisons should involve consideration of the potential contribution of bias to the p-value, confidence interval, or inference.





# ERDO – Evaluation Method

If RP denote the randomization procedure with  $P(\mathbf{z}) \neq \mathbf{0}$  for  $\mathbf{z} \in \{0,1\}^N$

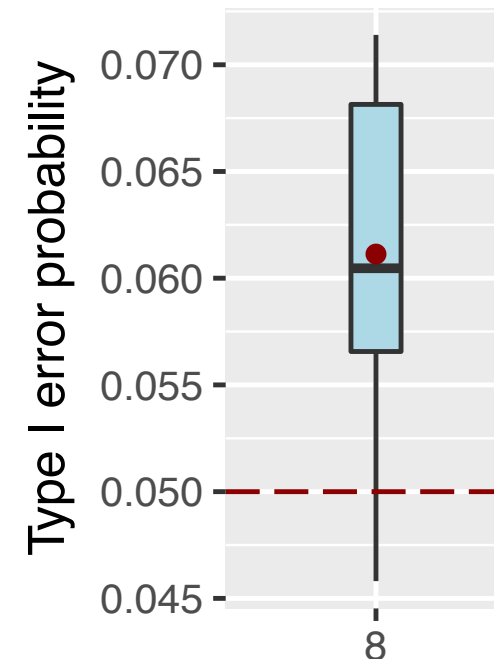
$$p_{RP,\tau}(\mathbf{z}) = F_{n-2k,-\delta(\mathbf{z},\tau),\lambda(\mathbf{z},\tau)}\left(t_{n-2k}\left(\frac{\alpha}{2}\right)\right) + F_{n-2k,\delta(\mathbf{z},\tau),\lambda(\mathbf{z},\tau)}\left(t_{n-2k}\left(\frac{\alpha}{2}\right)\right)$$

Mean Type 1 Error Probability

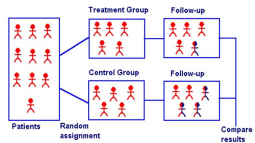
$$MTE = \sum_{\mathbf{z} \in \{0,1\}^N} p_{RP,\tau}(\mathbf{z}) P(\mathbf{z})$$

Go-No-Go criterion (actual Type I Error Probability)  
(Prob for preserve the 5% level)

$$GnG = \sum_{\mathbf{z} \in \{0,1\}^N} \mathbf{I}_{\{p_{RP,\tau}(\mathbf{z}) \leq \alpha\}} P(\mathbf{z})$$



**Actual Type I error probability when ignoring the bias in the analysis**



# Does Randomization the job?

## Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial

Craig M McDonald, Craig Campbell, Ricardo Erazo Torricelli, Richard S Finkel, Kevin M Flanigan, Nathalie Goemans, Peter Heydemann, Anna Kaminska, Janbernd Kirschner, Francesco Muntoni, Andrés Nascimento Osorio, Ulrike Schara, Thomas Sejersen, Perry B Shieh, H Lee Sweeney, Haluk Topaloglu, Már Tulinius, Juan J Vilchez, Thomas Voit, Brenda Wong, Gary Elfring, Hans Kroger, Xiaohui Luo, Joseph McIntosh, Tuyen Ong, Peter Riebling, Marcio Souza, Robert J Spiegel, Stuart W Peltz, Eugenio Mercuri, the Clinical Evaluator Training Group\*, and the ACT DMD Study Group\*

*Lancet, 2017*

### Summary

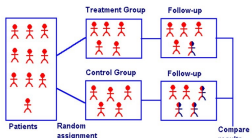
**Background** Duchenne muscular dystrophy (DMD) is a severe, progressive, and rare neuromuscular, X-linked recessive disease. Dystrophin deficiency is the underlying cause of disease; therefore, mutation-specific therapies aimed at restoring dystrophin protein production are being explored. We aimed to assess the efficacy and safety of ataluren in ambulatory boys with nonsense mutation DMD.

**Methods** We did this multicentre, randomised, double-blind, placebo-controlled, phase 3 trial at 54 sites in 18 countries located in North America, Europe, the Asia-Pacific region, and Latin America. Boys aged 7–16 years with nonsense mutation DMD and a baseline 6-minute walk distance (6MWD) of 150 m or more and 80% or less of the predicted normal value for age and height were randomly assigned (1:1), via permuted block randomisation (block size of four) using an interactive voice-response or web-response system, to receive ataluren orally three times daily (40 mg/kg per day) or matching placebo. Randomisation was stratified by age (<9 years vs ≥9 years), duration of previous corticosteroid use (6 months to <12 months vs ≥12 months), and baseline 6MWD (<350 m vs ≥350 m). Patients, parents and caregivers, investigational site personnel, PTC Therapeutics employees, and all other study personnel were masked to group allocation until after database lock. The primary endpoint was change in 6MWD from baseline to week 48. We additionally did a prespecified subgroup analysis of the primary endpoint, based on baseline 6MWD, which is reflective of anticipated rates of disease progression over 1 year. The primary analysis was by intention to treat. This study is registered with ClinicalTrials.gov, number NCT01826487.

Stratified PBR(4)

Put: 240 patients

# ERDO – Clinical Implementation Amount of Effect



Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial

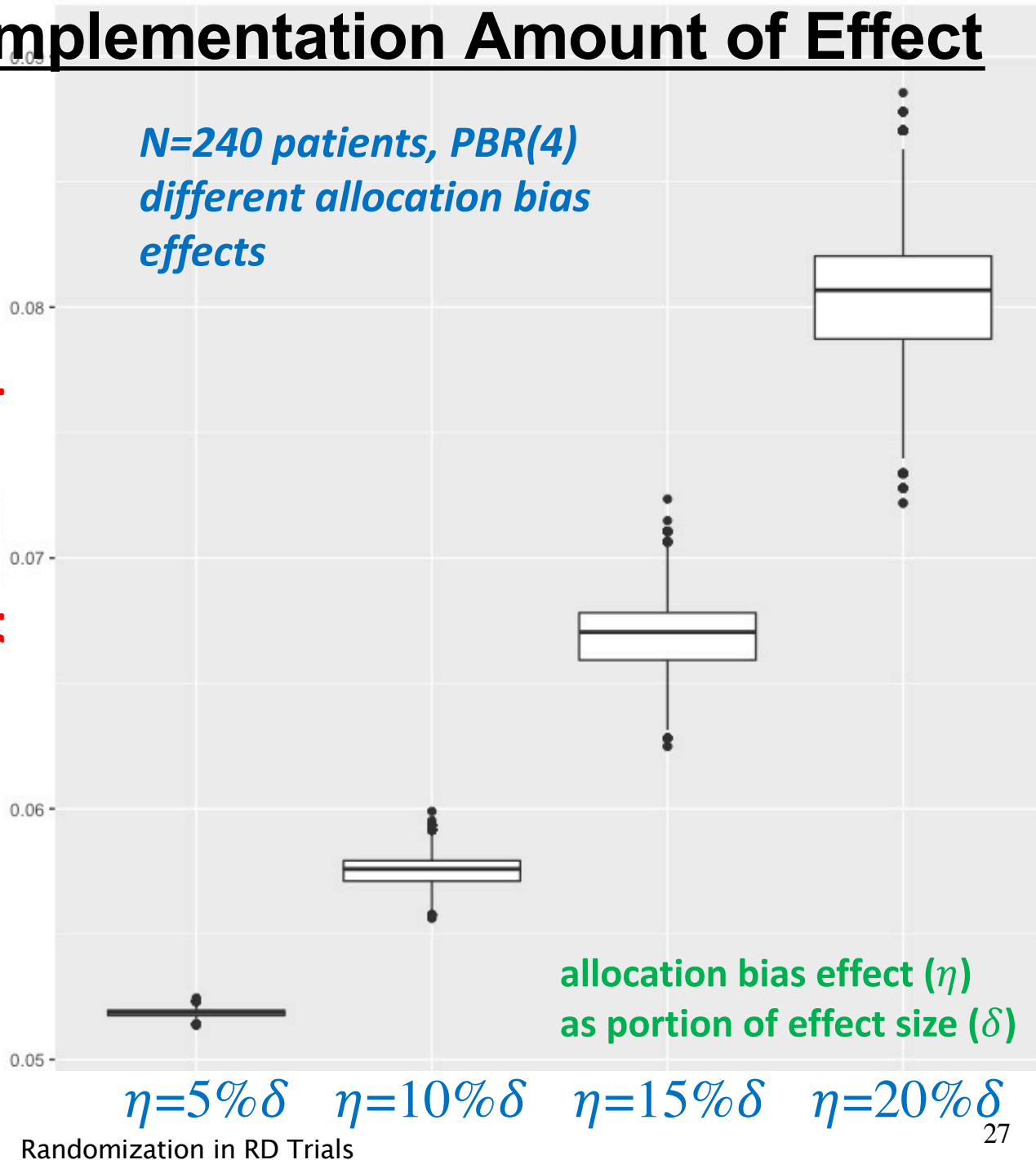
Craig M McDonald, Craig Campbell, Ricardo Erazo Torricelli, Richard S Finkel, Kevin M Flanigan, Nathalie Goemans, Peter Heydemann, Anna Kaminska, Janbernd Kirschner, Francesco Muntoni, Andrés Nascimento Osorio, Ulrike Schara, Thomas Sejersen, Perry B Shieh, H Lee Sweeney, Haluk Topaloglu, Már Tulinus, Juan Vilchez, Thomas Voit, Brenda Wong, Gary Elfring, Hans Kroger, Xiaohui Luo, Joseph McIntosh, Tuyen Ong, Peter Riebling, Marcio Souza, Robert J Spiegel, Stuart W Peltz, Eugenio Mercuri, the Clinical Evaluator Training Group\*, and the ACT DMD Study Group\*

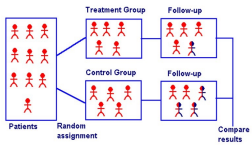
**Actual Type I error prob. increases with magnitude of allocation bias effect ( $\eta$ )**



Ralf-Dieter Hilgers

Actual Type I error prob.





# ERDO –Clinical Implementation Amount of Effect

Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial

*Craig M McDonald, Craig Campbell, Ricardo Erazo Torricelli, Richard S Finkel, Kevin M Flanigan, Nathalie Goemans, Peter Heydemann, Anna Kaminska, Janbernd Kirschner, Francesco Muntoni, Andrés Nascimento Osorio, Ulrike Schara, Thomas Sejersen, Perry B Shieh, H Lee Sweeney, Haluk Topaloglu, Mår Tullinius, Juan Vilchez, Thomas Voit, Brenda Wong, Gary Elfring, Hans Kroger, Xiaohui Luo, Joseph McIntosh, Tuyen Ong, Peter Riebling, Marcio Souza, Robert J Spiegel, Stuart W Peltz, Eugenio Mercuri, the Clinical Evaluator Training Group\*, and the ACT DMD Study Group\**

What can be done better?

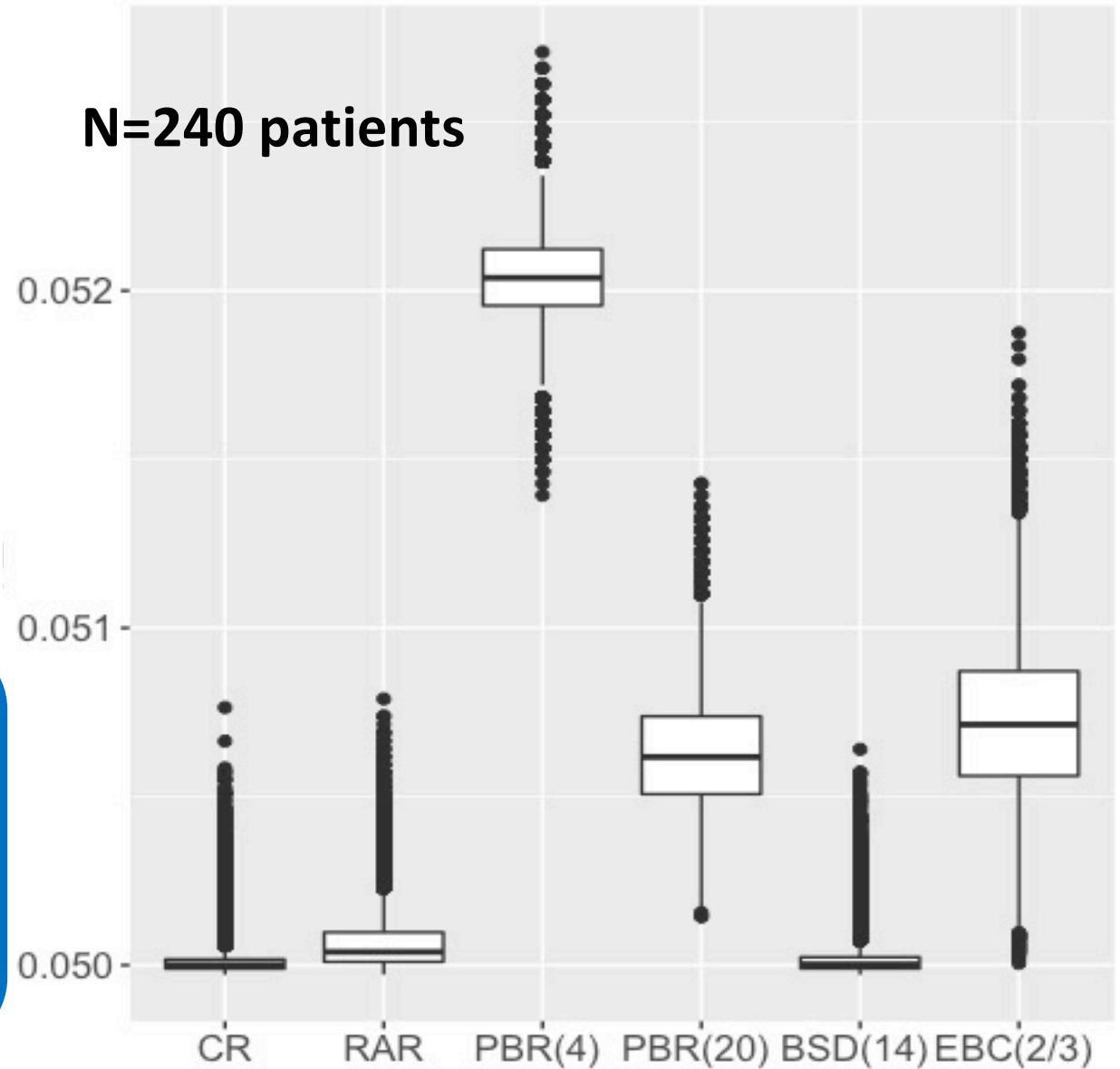
Comparison of procedures

GnG criterion seem to be more sensitive

Large difference between procedures

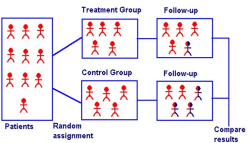


N=240 patients



GnG=44% GnG=12% GnG=0% GnG=0% GnG=39% GnG=0%  
 $\overline{sig} = 0.05$   $\overline{sig} = 0.05$   $\overline{sig} = 0.05$   $\overline{sig} = 0.05$   $\overline{sig} = 0.05$   $\overline{sig} = 0.05$

Randomization in RD Trials



# ERDO – Discussion and Clinical Implementation

**Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial**

*Craig M McDonald, Craig Campbell, Ricardo Erazo Torricelli, Richard S Finkel, Kevin M Flanigan, Nathalie Goemans, Peter Heydemann, Anna Kaminska, Janbernd Kirschner, Francesco Muntoni, Andrés Nascimento Osorio, Ulrike Schara, Thomas Sejersen, Perry B Shieh, H Lee Sweeney, Haluk Topaloglu, Már Tullinius, Juan J Vilchez, Thomas Voit, Brenda Wang, Gary Elfving, Hans Kroger, Xiaohui Luo, Joseph McIntosh, Tuyen Ong, Peter Riebling, Marcio Souza, Robert J Spiegel, Stuart W Peltz, Eugenio Mercuri, the Clinical Evaluator Training Group\*, and the ACT DMD Study Group\**

sb	Placebo	Atalure
-1	58	6
0	44	52
1	8	52
	<b>110</b>	<b>110</b>

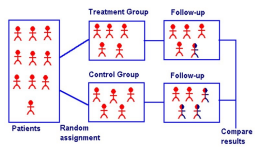
Estimated Effect from simulated data with PBR(4) (without stratification) N=220

Scenario	Atalure Mean (SD)	Placebo Mean (SD)	Unadjusted	Adjusted Analysis		Estimated Bias
			t-Test	Treatment effect	Bias Effect	
1	38.16 (99.25)	13.83 (99.97)	0.0716	0.0415	0.4573	1: -1.89 0: 0.00 -1: -22.04
2	73.48 (109.72)	101.39 (98.87)	0.0487	0.1515	0.9502	1: 5.13 0: 0.00 -1: -1.63

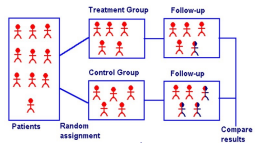
**Main Analysis**

**Sensitivity Analysis**

**Future Planning**



# WHICH DESIGNS ARE DEVELOPED RIGHT NOW?



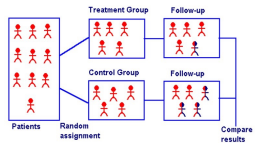
# Which Design Settings are developed?

I assume there is something for you here



Design	Reference
2-Arm, continuous PeV	Hilgers, ERDO, 2017
<i>2-Arm unbalanced allocation, continuous PeV</i>	<i>Hilgers manuskript [EJP-RD]</i>
<i>2-Arm group sequential, continuous PeV</i>	<i>Mullenmeister [EJP-RD]</i>
<i>2-Arm, multiple PeV</i>	<i>Schoenen [iSTORE ]</i>
<i>2-Arm, slope model based on LMEM</i>	
<i>2-Arm, binary PeV</i>	<i>Reugels</i>
Multicenter center 2-Arm, continuous PeV	Hilgers, SMMR, 2019
2-Arm unbalanced, time to event PeV	Rückbeil 2017,19
2-Arm unbalanced, time to event PeV, delayed event	Rückbeil 2021
Multiarm, continuous PeV	Uschner, 2018

*What is about Platform Trials -> of course we are working on this.*

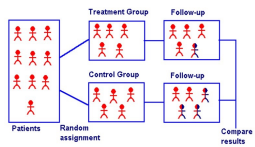


# Some key facts from our investigation?

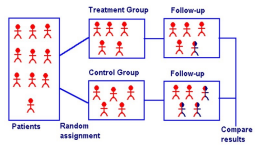
Design	(Type I Error probability ) is...
2-Arm, continuous PeV	.... elevated with larger studies [PBR]
<i>2-Arm unbalanced allocation, continuous PeV</i>	.... elevated with knowledge about allocation ratio
<i>2-Arm group sequential, continuous PeV</i>	.... elevated with larger number of stages .... Fairly similar with Pocock's and O'Brien's design <i>(depending on balancing)</i>
Multicenter center 2-Arm, continuous PeV	.... similar with Strat. & unstrat. Randomization ... similar with unbalanced center sample sizes .... elevated with larger number of centers
2-Arm unbalanced, survival PeV	.... affected by the randomization procedure
2-Arm unbalanced, survival PeV, delayed event	.... affected by test statistic, type and strength of allocation bias and the randomization procedure
Multiarm, continuous PeV	.... elevated with more groups [PBR]





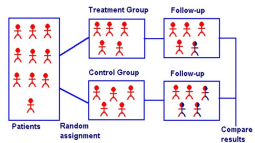


# WHAT'S ABOUT REAL APPLICATION?



# Nicofa Trial

Lesson learned	References
<p>Progression of SARA via LMEM</p> <ul style="list-style-type: none"> <li>-&gt; Model Parameters</li> <li>-&gt; Effect Estimate</li> <li>-&gt; Selection Bias effects</li> </ul>	<p>Reetz et al. <a href="#">Progression</a> characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 2 year cohort study Lancet Neurol 2016 doi: 10.1016/S1474-4422(16)30287-3.</p> <p>Reetz et al. <a href="#">Progression</a> characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 4-year cohort study Lancet Neurol 2021 doi: 10.1016/S1474-4422(21)00027</p>
<ul style="list-style-type: none"> <li>-&gt; amount of time trend and selection Bias effects</li> <li>-&gt; estimated enrollment / center</li> <li>-&gt; valuation criterion</li> </ul>	<p>Reetz et al. <a href="#">Protocol</a> of a randomized, double-blind, placebo-controlled, parallel-group, multicentre study of the efficacy and safety of nicotinamide in patients with Friedreich ataxia (<a href="#">NICOFA</a>) Neurological Research and Practice 2019 <a href="https://doi.org/10.1186/s42466-019-0038-9">https://doi.org/10.1186/s42466-019-0038-9</a> ClinicalTrials.gov Identifier: NCT03761511</p>



# Nicofa Trial

## Lesson learned

- Progression of SARA via LMEM
- > Model Parameters
- > Effect Estimate
- > Selection Bias effects

## References

Reetz et al. [Progression](#) characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 2 year cohort study Lancet Neurol 2016 doi: 10.1016/S1474-4422(16)30287-3.

Reetz et al. [Progression](#) characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 4-year cohort study Lancet Neurol 2021 doi: 10.1016/S1474-4422(21)00027

### Randomization and blinding

The randomization list prepared by the Department of Medical Statistics of the RWTH Aachen University Hospital, Aachen in Germany using randomizeR is stratified by center. The best practice randomization procedure to minimize the impact of selection and time trend bias on the type one error will be selected via a simulation study (ERDO) [19]. The packaging of the investigational product and placebo following the randomization list will be done in Mainz, Germany, labelled with a randomisation code. This will maintain concealment and double blinded treatment allocation. After randomization neither the patients nor the investigator or sponsor will be aware of the treatment allocation. Patients assigned to one of the double-blinded treatments will take nicotinamide capsules or matching placebo. The capsules will be identical in

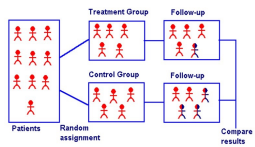
Ralf-Dieter Hilgers

etz et al. [Protocol](#) of a randomized, double-blind, placebo-controlled, parallel-group, multicentre study of the efficacy and safety of nicotinamide in patients with Friedreich ataxia (NICOFA) Neurological Research and Practice 2019 <https://doi.org/10.1186/s42466-019-0038-9>

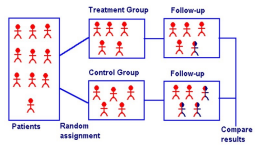
*ClinicalTrials.gov Identifier: NCT03761511*

*E-RARE Program (ERARE16-FP-045)*

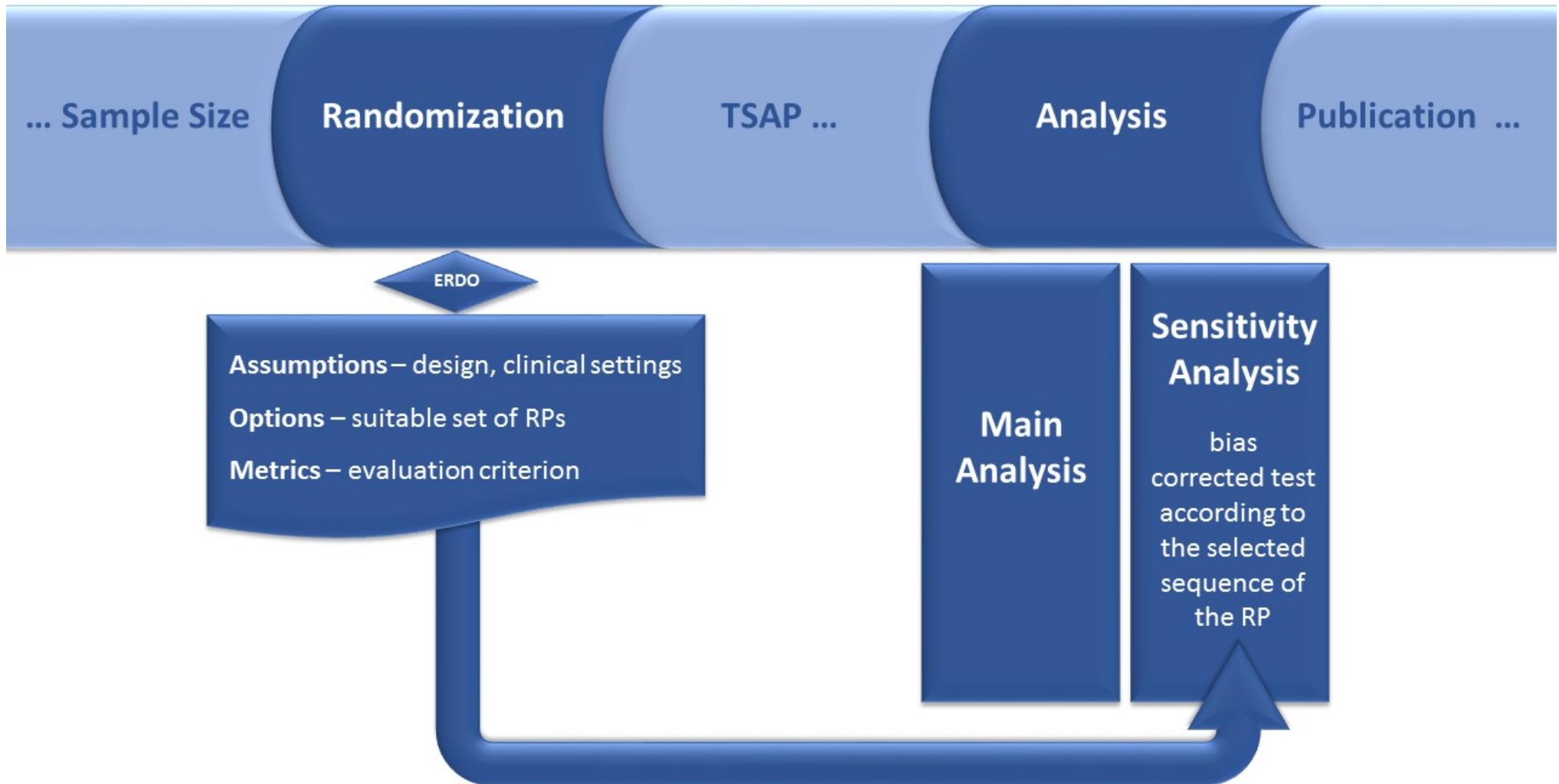
*DFG SCHU 932/10-1, KL 795/4-1*

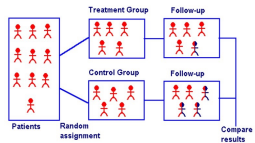


# SUMMARY



# ERDO in the Analysis





# The journey continues

## *EpiSTOP – IDeAI*

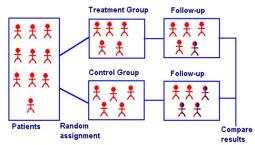
- Assess the level of evidence link between the treatment allocation process and the analysis of the primary endpoint variable is used to quantify the impact of the level of evidence and by this quantify the uncertainty of trial result

## *iSTORE*

- Develop randomization-based models as alternative analysis strategies and explore the level of evidence
- Bias models for assessment with multiple endpoints
- Randomization Based Inference with multiple endpoints

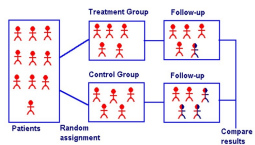
## *EvidenceRND*

- develop a randomization-based inference framework enabling valid and efficient RCTs in ultra-rare RNDs, allowing to evaluate for (i) multiple biases and (ii) missing data, and to explore the resulting level of evidence.



# Summary

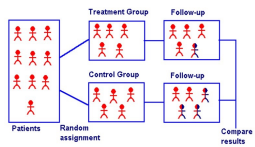
- Randomization is important in assessing clinical trial validity
- Randomization and selection of the „best performing procedure“ is even more important in RD trials
- ERDO provides a useful approach & combined with the randomizeR an efficient tool for improving RD clinical trials
- Estimate of Bias can now be used in Evidence Synthesis / HTA
  - Topic in EPISTOP-IDeAI
- Within the EJP-RD Innovation projects iSTORE and EVIDENCE-RND we will use randomization to assess level of evidence in finite (limited) population RD's

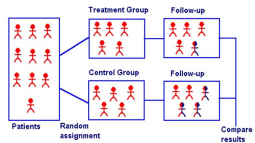


# References

- Kennes LN, Cramer E, Hilgers RD and Heussen N. (2011). The impact of selection bias on test decisions in randomized clinical trials SIM 2011; DOI: 10.1002/sim.4279.
- Tamm M, Cramer E, Kennes LN and Heussen N. Influence of Selection Bias on the Test Decision - A Simulation Study Methods of Information in Medicine 2012, DOI: 10.3414/ME11-01-0043.
- Tamm M and Hilgers RD. Chronological Bias in Randomized Clinical Trials Arising from Different Types of Unobserved Time Trends Methods of Information in Medicine 2014; DOI: 10.3414/ME14-01-0048.
- Rückbeil M, Hilgers RD and Heussen N. Assessing the impact of selection bias on test decisions in trials with a time-to-event outcome SIM 2017 DOI: 10.1002/sim.7299
- Hilgers RD, Uschner D, Rosenberger WF, Heussen N. ERDO - a framework to select an appropriate randomization procedure for clinical trials. BMC Medical Research Methodology 2017 DOI: 10.1186/s12874-017-0428-z
- Rückbeil M, Hilgers RD and Heussen N. The impact of selection bias in randomized multi-arm parallel group clinical trials. PlosOne 2018 DOI: 10.1371/journal.pone.0192065
- Uschner D, Schindler D, Heussen N and Hilgers RD. randomizeR: An R Package for the Assessment and Implementation of Randomization in Clinical Trials J Statistical Software 2018
- Hilgers RD, Manolov M., Rosenberger WF, Heussen N. Design and analysis of stratified clinical trials in the presence of bias Statistical Methods in Medical Research 2019; doi.org/10.1177/0962280219846146 |: .

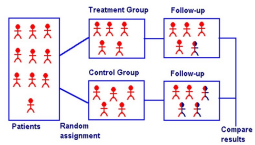






# Spring and Autumn Webinar Series

- dedicated to inform about innovative trial methodology tailored to RD clinical trials
- Topics related but not limited to **EJP-RD Demonstration Projects & Innovation Projects**
- Please take a look to EJP-RD website for registration
- Past participants will receive an announcement



announcement



## 2nd Spring-Webinar 6th May 2022

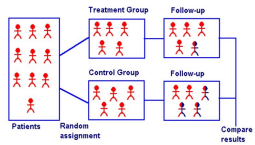
by Johan Verbeeck



# Topic: Composite endpoints including patient relevant endpoints (Quality of Life)

EpiStopIDeAI

This research is part of the EU-FP7 IDeAI project (GA No. 602552)



Pre-announcement



## 3rd Autumn Webinar November 2022

# Topic: External and historical data use in clinical trials

This research is part of the EU-FP7 IDeAl project (GA No. 602552)