Explorative klinische Versuche

Was ist das und warum braucht es das?

Sven Trelle, CTU Bern

CTU Lecture, 26.10.2022

The plan for today

3 broad topics

– Background

– Single-arm exploratory trials

– Randomized exploratory trials

Disclaimer and conflicts of interest
with regard to this lecture

- I am Director of an academic clinical trials unit (CTU Bern)

- Anything that promotes patient-oriented clinical research or makes it more attractive profits us/me directly or indirectly …

- The more complicated and sophisticated the methods I present, the more likely a specialist (at CTU Bern) needs to be involved …
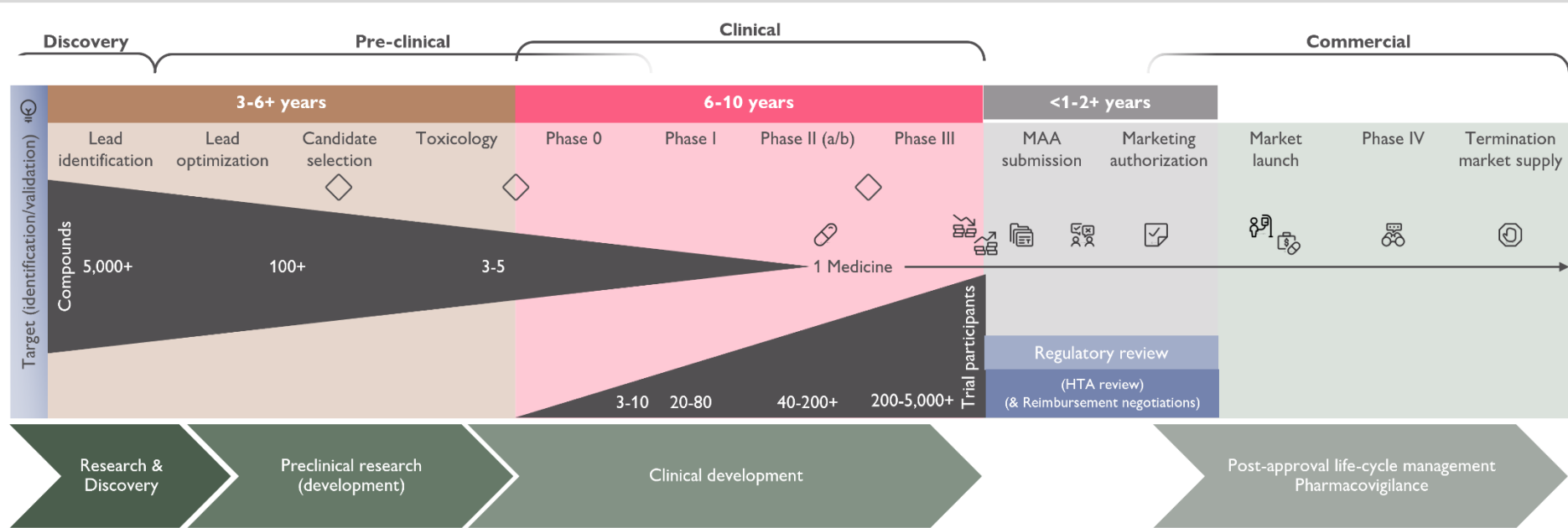
Trial designs
Cave!

– Most trial designs have been developed for cancer drugs (traditional/chemotherapeutic and targeted)

– Most of the examples I present are from oncology

– A lot of concepts are broadly applicable to other disease areas and also other health-related interventions
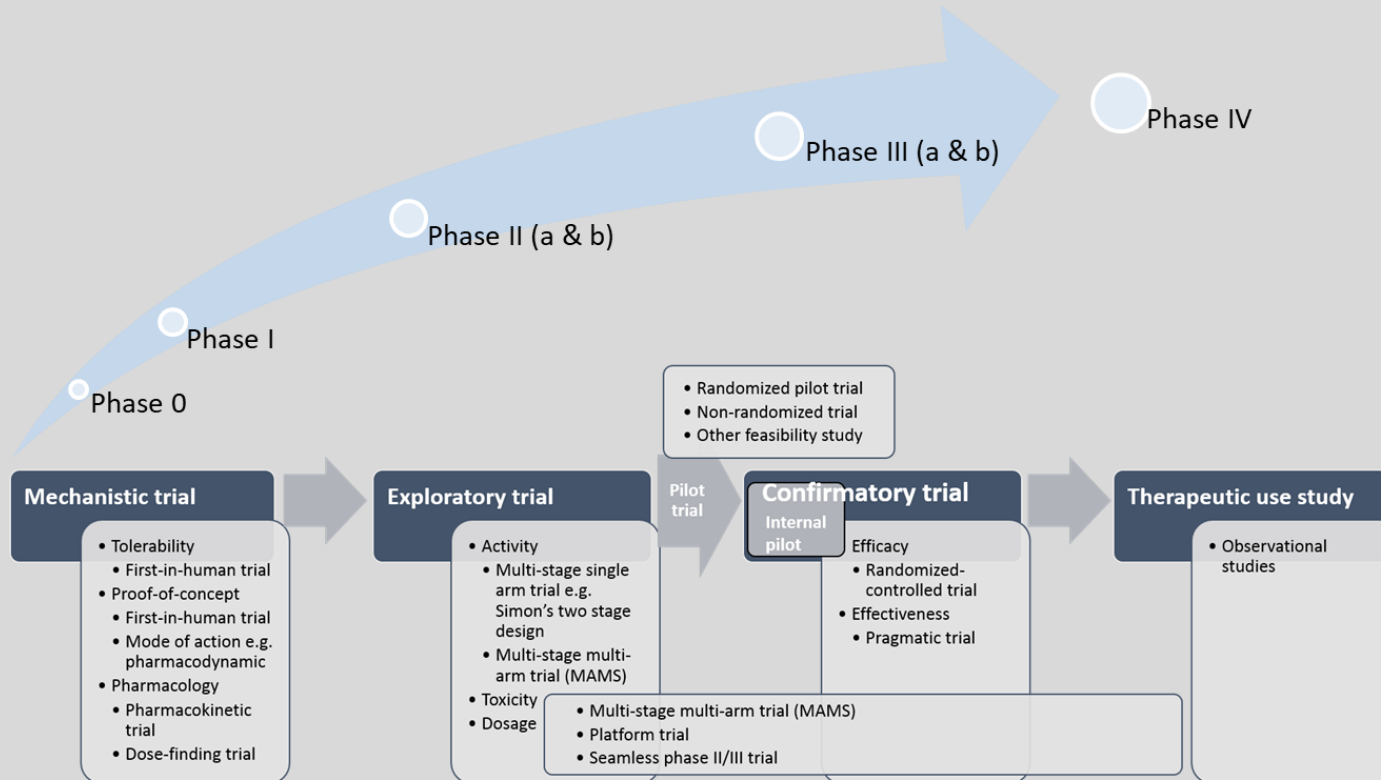
Drug development

From research to market launch and beyond

Clinical drug development
Phases and related trials

Exploratory clinical trials
Main objective

- To support the decision on whether to conduct a confirmatory trial (*is it worth to invest in further trials i.e., larger randomized trials?*)
  - Has the intervention sufficient/promising activity?
  - To identify the most promising dose, schedule etc.
  - To get a better idea on toxicity.

  - …

- **Not** whether intervention is worth to be used in practice (→ confirmatory trial(s))

→ *Gatekeeper/Proof of concept*

Implications

Exploratory trials as gatekeeper

– Trial design: single-arm or non-comparative randomized trial (this talk)

– Study population: maybe more selected

– Primary endpoint: clinical relevance$\downarrow$ (activity not efficacy/effectiveness!)

– Trial length: relatively short term

– Type I error: maybe more liberal (instead of 0.05 (two-sided) maybe 0.10 or even 0.20?)

– Power: maybe more stringent (e.g., 90 or 95% instead of 80%)

– Conclusions: considering the above …

Phased approach to drug development
Why (beyond ethics)?

Dowden & Munro 2019

U.S.A Election

I usually always win!
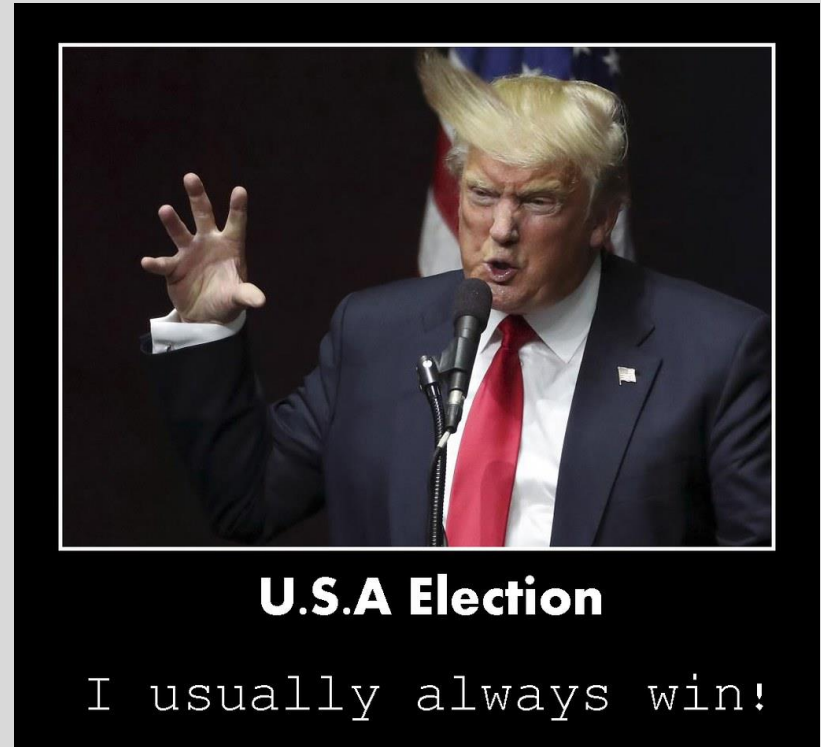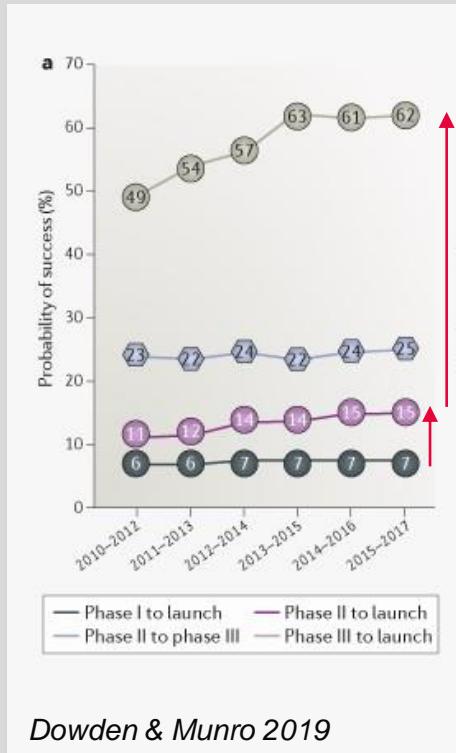
Overarching objectives of a confirmatory trial
To set the scene

- Experiment to quantify cause-and-effect i.e. exposure/intervention → outcome

- Mechanistic (scientific research)

- (Clinical) Practice (evaluative research)
  - Commercial/industry: to sell a product (e.g. pharmaceutical, device, …) to make money
  - Academic: to change practice, make a career, …
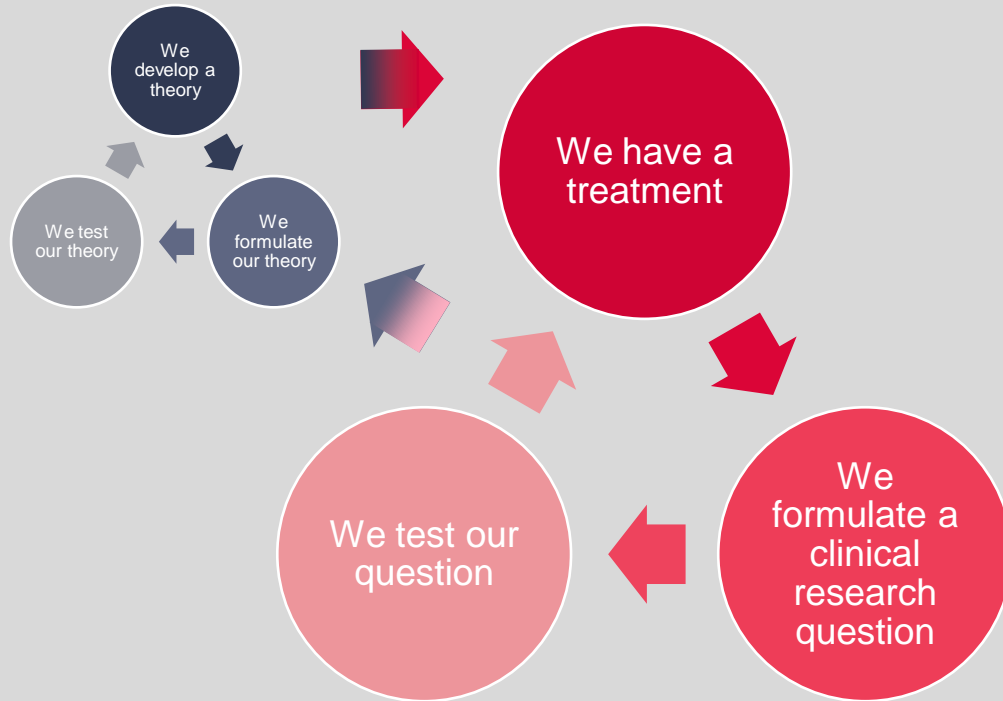 - Mandated (UK NIH): to resolve uncertainty and optimize health care provision

Discovery (via proofs)

Objective of (clinical) experimenting
# Proofing

Exploratory trials

Navigational decision-making

The exploratory trial landscape
Academic (and industry)

- 1990-96: "… quality of the statistical component … is generally poor …" (Mariani & Marubini 2000)

- 1995 → 2000: "… one-half of articles are still published despite the fact that they do not, or hardly, mention the statistical method employed." (Thezenas et al. 2004)

- 2011: "Primary end point definition, justification of sample size and definition of the evaluable population were reported in only 107 (68.6%), 121 (77.6%), and 52 (33.3%) cases, respectively." (Grellety et al. 2014)

- 2005 → 2014: "We found that the proportion of trials where the null hypothesis was formally tested increased in 2010 and 2014 compared to 2005." (Ivanova et al. 2016)

The most often used exploratory design
Simon's two stage design (>2,700 cited)

Web of Science (Oct 2022)
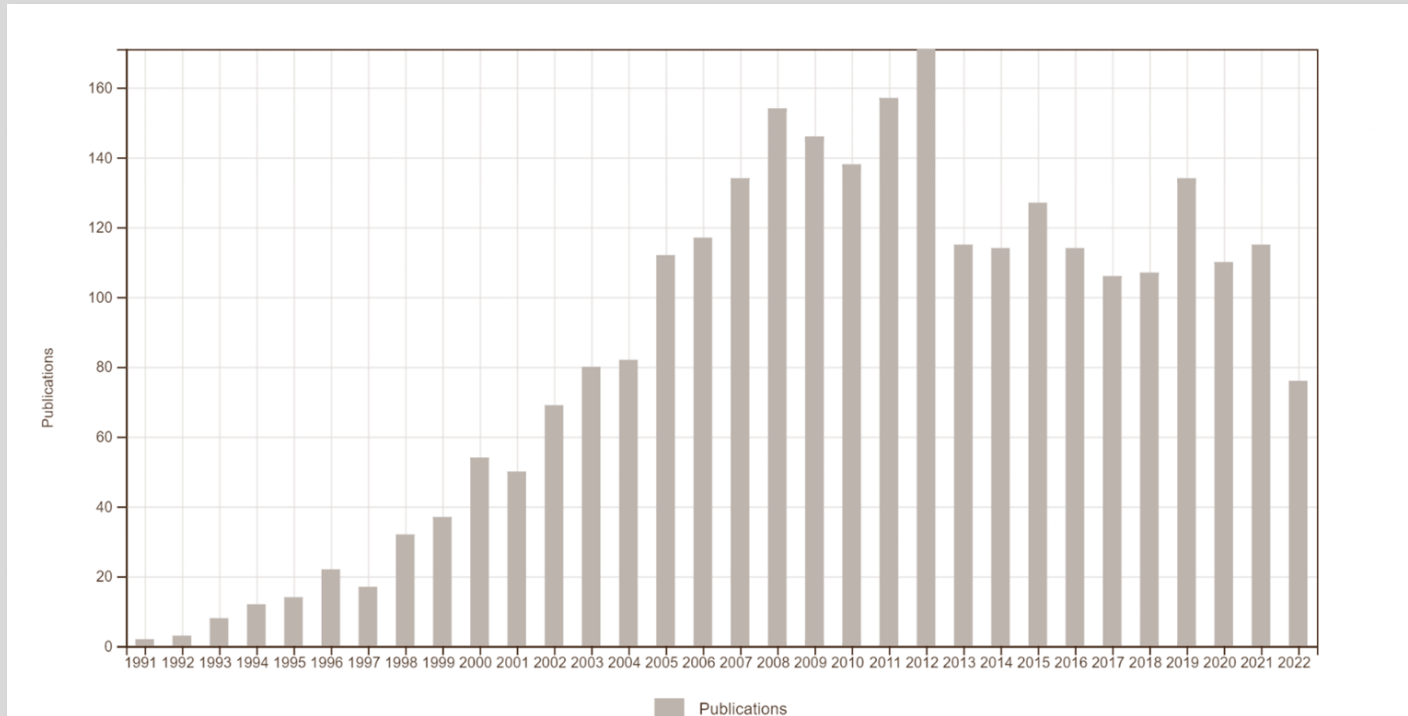
Medical disciplines

Simon's two stage design

Web of Science (Oct 2022)

The principle
Simon's two-stage design

- Usually binary outcome (proportion) e.g. response (regression of tumor)
- Fix a target (promising) proportion
- Fix a proportion that is uninteresting (null outcome)
- Fix type I (one-sided; alpha) and type II (power) error
  - Alpha higher than 0.05 desirable
  - Power higher than 0.80 desirable

- Two stages/phases with a single-arm design
- Decision rule: trial success yes/no

- After first stage: test for futility (trial stops if chances to observe the target proportion are too small)

Simon 1989

Proof-of-concept
Simon's two-stage design

– Usually binary outcome (proportion) e.g. response

– Fix a target (promising) proportion

– Fix a proportion that is uninteresting (null outcome)

– Fix type I (alpha) and type II (power) error

– Two stages/phases with a single-arm design

– Decision rule: trial success yes/no

– After first stage: test for futility (trial stops if chances to observe the target proportion are too small)

– Closely related to non-inferiority trials

– Implicitly historically controlled trial (target and null response)

Principle objective

Simon's two stage design

– To determine whether a treatment is active (not efficacious and surely not whether it is effective)

– The concept of 'trial success'

→ The hypothesis testing strategy is key to understand the design and underlying objective

Variants

Modifications of Simon's design

- Optimal: assuming null outcome, minimize sample size of first stage (expected sample size)
- Minimax: Minimize overall sample size

- Two simultaneous binary outcomes e.g. response and toxicity (Bryant & Day 1995)
- Three or more stages (Ensign et al. 1994; Chen 1997; Chen et al. 1994)
- Continous outcome (Wason et al. 2011)
- Ordinal outcome (Ivanova et al. 2012)
- Intermediate outcome for first stage (Kunz et al. 2015)
- Early stop not only for futility but also for efficacy (Mander & Thompson 2010)
- Bayesian framework (predictive probability, more flexible, allows to incorporate uncertainty in historical data; not better if target is (very) low; Liu et al. 2017)
- …

TUXEDO-1

First example

- To evaluate efficacy and safety of trastuzumab deruxtecan [antibody–drug conjugates (ADCs)] in … HER2-positive breast cancer patients with active brain metastases … and more broadly as proof of principle for the intracranial activity of ADCs.
  - Primary endpoint: best response of brain metastases (after >/= 1 cycle)
  - Promising: response rate >/=61%
  - Uninteresting: response rate </=25%
  - Alpha 5%, Power 80%
- 6 + 9 (N=15) patients
- For trial success: >/=7/15 (>/=3/6) responses

Bartsch 2022

Results
TUXEDO-1

- Overall, 15 patients enrolled over 1 year

- After stage I (N=6)
  - 5 (83%) responded

- After stage II (N=15)
  - 11 (73%) responded (2 complete remission, 9 partial)
  - →Criteria for trial success met

- →Clear conclusions

- →No post-hoc, data-driven decisions

Bartsch 2022

NCT00148018

Bortezomib+dexamethasone in Hodgkin's lymphoma

– Primary endpoint: best response

– Promising: response rate >20%

– Uninteresting: response rate <5%

– Alpha 10%, Power 90%

– 12 + 25 (N=37) patients

– For trial success: >/=4/37 (>/=1/12) responses

– Overall, 12 patients enrolled over 13 months

– No responses observed after 12 patients → trial stopped

Trelle 2007

No concurrent comparator
Main issue

- Target fixed from historical data

- Separating treatment effect from trial effects (current trial vs. historical data)
  - Patient selection
  - Co-interventions, supportive care
  - Other time effects, confounders, biases

- Degree of problem depends on outcome e.g.,
  - spontaneous or variable outcomes
  - tumor response might be less prone than progression-free or overall survival
  - …

See for example Rubinstein 2009

Randomized exploratory trial designs
Not confirmatory!

1. Non-comparative approaches where randomized trial arms are considered separately as if they were single-arm trials

2. Comparative approaches where
   a. Several experimental treatments are compared to select the most promising one for future trials and
   b. Experimental treatment(s) are compared to standard of care (or placebo) to screen whether an experimental treatment is actual worth for evaluating it in a definitive confirmatory trial.

– MultiArm MultiStage trials (MAMS): 2a and 2b
– Platform trials

Randomized non-comparative
FCM (standard) vs. FCM-R in CLL

- A randomized phase II trial of fludarabine, cyclophosphamide and mitoxantrone (FCM) with or without rituximab in previously treated chronic lymphocytic leukemia
    - Randomization (1:1) to 6 cycles every 28 d: Fludarabine, Cyclophosphamide, Mitoxantrone +/- Rituximab (*to protect against a possible patient selection bias, and give internal and external validity of the results*)
    - Gehan's two-stage design
        - Promising: ≥40% overall response rate
        - Alpha 0.1, power 90%
        - 4+21 (N=25) patients
    - 2x26 patients planned, stage I after 4 patients in FCM-R arm only (3/4 patients responded → continue)

- Overall, 52 patients enrolled over 19 months
- → *The overall response rates to FCM and FCM-R were 58% and 65% respectively.*

Hillmen 2011

Randomized comparative
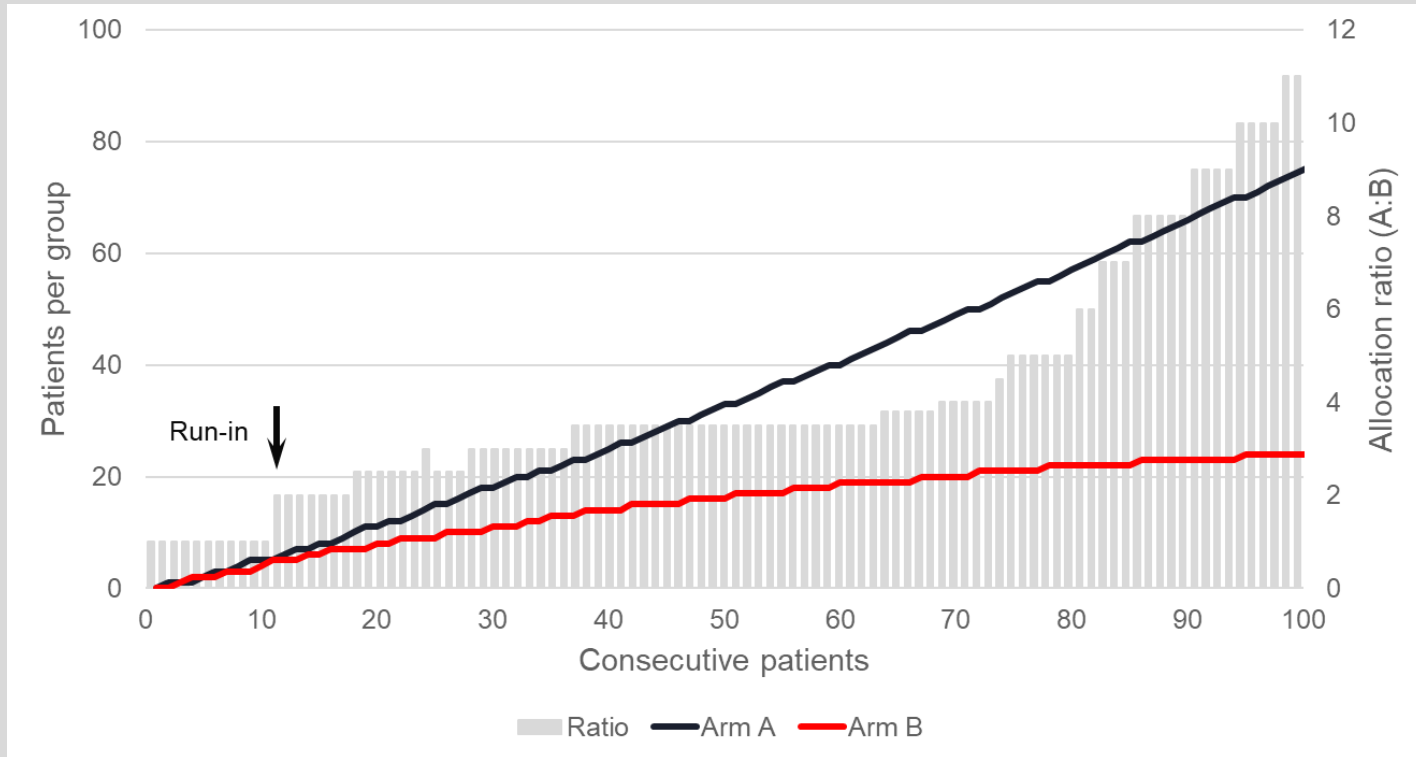## Several experimental

- A randomized phase II study of standard-dose versus high-dose rituximab with BEAM in autologous stem cell transplantation for relapsed aggressive B-cell non-hodgkin lymphomas

- Bayesian adaptive design
  - Comparison of disease-free survival between the 2 arms is constantly updated based on accumulating data
  - Allocation was adapted for each patient (based on comparison i.e. higher chances for better arm)
  - Sample size 100 (probably feasibility)

Srour 2017

Adaptive randomization
Versus fixed allocation ratio

– Fixed allocation ratio: all participants have the **same** probabilities to be allocated to the different arms **over the whole** trial period

– Most often: 1-to-1 i.e. 50% chance

– Alternatives
  – 2:1 i.e. two times higher chance to be allocated in one arm versus the other
  – 1:1:1:1:… i.e. multi-arm trial but probabilities are similar across arms
  – N:1 …

– Adaptive randomization → allocation ratio changes over the course of the trial

Adaptive randomization
Possible scenario with one better arm

Randomized comparative

Several experimental

– A randomized phase II study of standard-dose versus high-dose rituximab with BEAM in autologous stem cell transplantation for relapsed aggressive B-cell non-hodgkin lymphomas

– Bayesian adaptive design

  – Comparison of disease-free survival between the 2 arms is constantly updated based on accumulating data

  – Allocation was adapted for each patient (based on comparison i.e. higher chances for better arm)

  – Sample size 100 (probably feasibility)

→ " … after 93 patients [stopped] because it was considered extremely unlikely that either treatment arm would have statistically superior DFS should the trial enroll the full complement of the originally planned 100 patients."
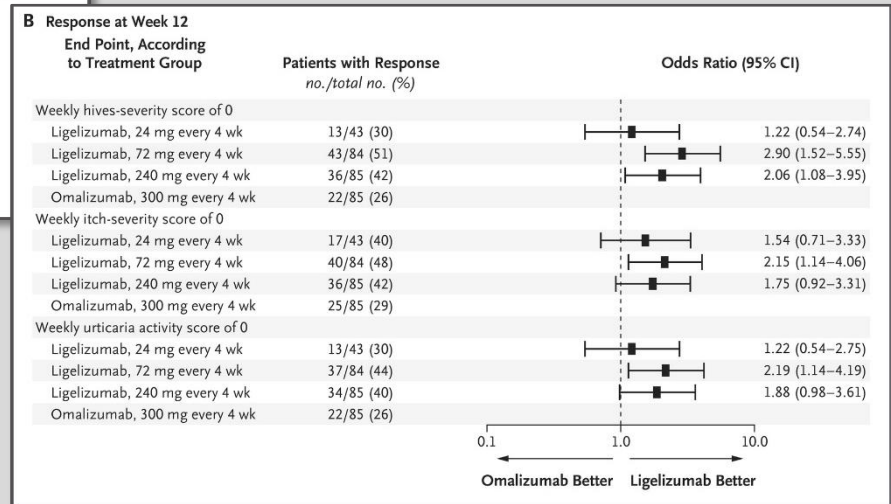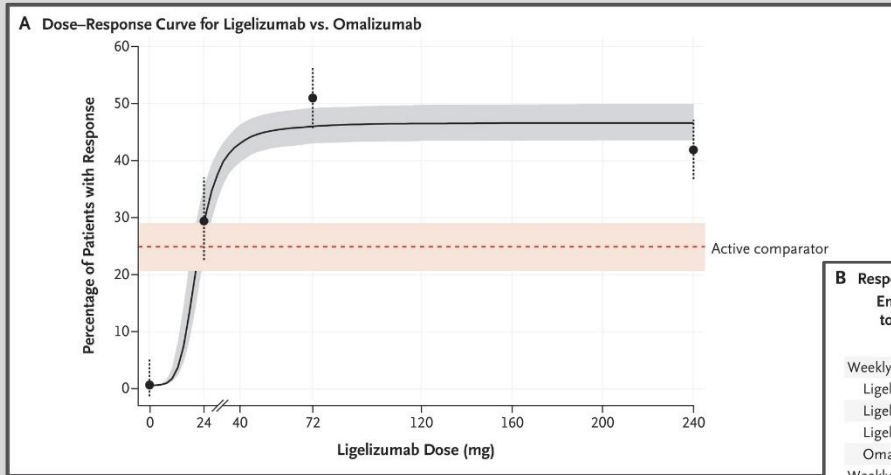
Srour 2017

Randomized comparative

Several experimental → Dose determination

u^b

b
**UNIVERSITÄT
BERN**

– Ligelizumab for chronic spontaneous urticaria

– Primary endpoint: complete control of hives ('response') at 12 weeks
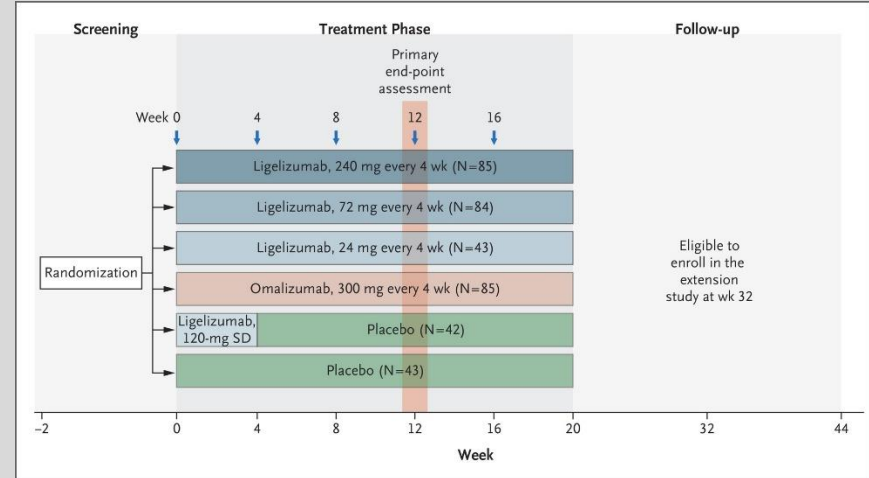
Maurer 2019

Ligelizumab trial
Maurer 2019

Single-arm versus randomized designs

"Why randomizing?"

– Historical data versus sample size

– Efficiency of design critically depends on accuracy and precision of historical data/estimate (Pond & Abbasi 2011, Sambucini 2015)

Take home message

– Not confirmatory but to decide whether confirmatory trial worthwhile

– Endpoint and study length less clinically relevant

– Formalized study design
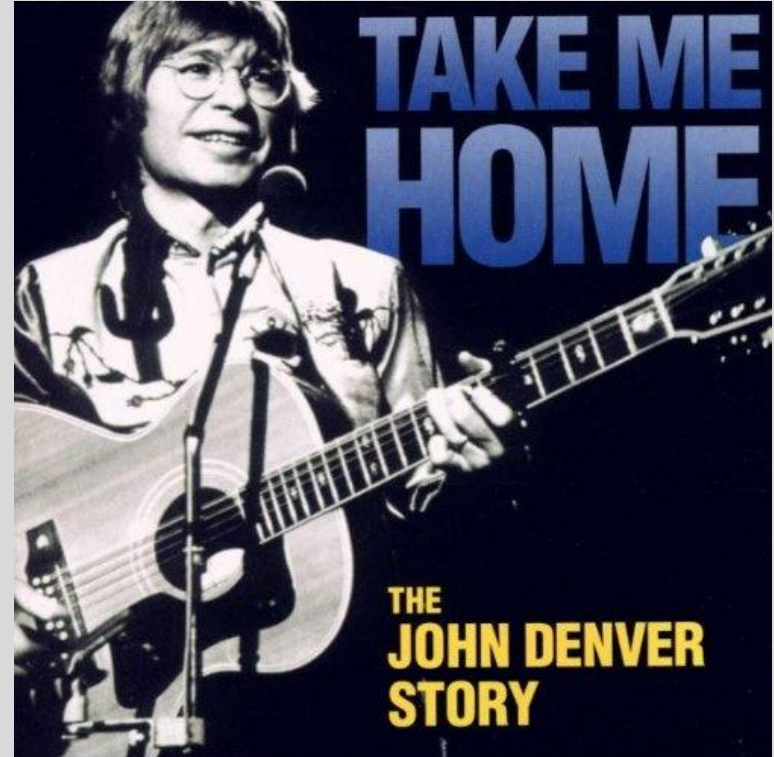
Thank you

for your attention

Sven Trelle, CTU Bern

u^b

UNIVERSITÄT
BERN

References
- Bartsch R et al. Nat Med 2022; 28: 1840-7.
- Brown SR et al. Br J Cancer 2011; 105: 194-9.
- Dowden H. Nat Rev Drug Discov 2019; 18: 495-6.
- Grellety T. Ann Oncol 2014. 25: 536-41.
- Hillmen P et al. Br J Haematol 2011; 152: 570-8.
- Ivanova A et al. J Biopharm Stat 2016; 26: 141-9.
- Mariani L, Marubini E. J Clin Oncol 2000; 18: 429.
- Maurer M et al. N Engl J Med 2019; 381: 1321-32.
- Pond GR, Abbasi S. Clin Trials 2011; 8: 260-9.
- Paoletti X et al. Ann Oncol 2015; 26: 1808-12.
- Rubinstein L et al. Clin Cancer Res 2009; 15: 1883-90.
- Sambucini V. J Biopharm Stat 2015; 25: 474-89.
- Simon R. Control Clin Trials 1989. 10: 1-10.
- Srour AS et al. Br J Haematol 2017. 178: 561-70.
- Thezenas S et al. Eur J Cancer 2004; 40: 1244-9.
- Trelle S et al. Haematologica 2007; 92: 568-9.
Pictures
- Stairway by Free-Photos from Pixabay