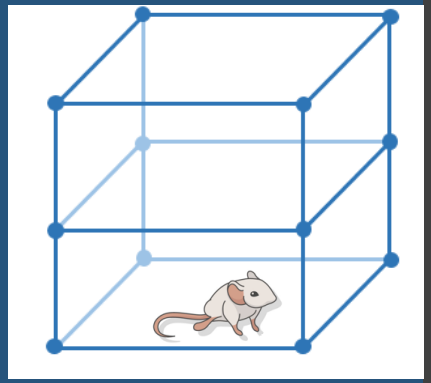
Many factors, few prior information –

some thoughts, approaches and examples to consider for sample size calculation in preclinical animal trials



Dr. Andreas Allgöwer

Institute for Epidemiology and Medical Biometry

Picture from Freepik

Some thoughts and approaches about many factors…

- factors: sex, genotype, time points, treatment, cage,…
  - fixed, random effects
  - crossed, nested

- when setting up a new animal model:
  - common approach: one factor at a time not very efficient
  - large factorial designs more efficient
- when generating hypotheses:
  - small factorial designs
  - block designs
- when testing hypotheses a stricter experimental procedure should be done.
  - block designs
  - planned comparisons or other multiple comparison procedures

Preclinical animal experiment → biometric report

10.1. Versuchstyp

A	☐	Technisch erforderliche (z.B. die Verwendung von Tieren zur Gewinnung von Material)
B	☐	Hypothesen generierender Versuch **mit geringen Tierzahlen**, bei dem noch keine spezifizierten Hypothesen geprüft werden können (Pilot-Versuch, Grundlagenklärung)
C	☒	Hypothesen überprüfender Versuch → **in diesem Fall je ein Formblatt zur biometrischen Planung pro Teilversuch beifügen** (alternativ biometrisches Gutachten)
D	☐	Sonstige nicht-hypothesengenerierende Versuche (z.B. Versuchsvorhaben zu Aus-, Fort- und Weiterbildungszwecken)

in english

10.1. experiment type

A technically necessary (e.g gaining material)

B hypothesis generating experiment with fewer animals

C hypothesis testing experiment with one biometric planning form per sub-sample (alternative biometric report)

D other non-hypothesis generating experiments (e.g. training, educational purposes)

Challenges in planning and analysis of preclinical animal experiments
(sample size calculation perspective)

- small sample sizes

  → parametric vs. non-parametric analysis

- sparse prior data

  → statistical test for the actual animal experiment may not be applicable for sample size calculation

- mistaken statistical test- and result-thinking („other research groups do it the same way“)

  → inaccurate experimental design

→ limited validity of calculated sample size

> “In principle, the true sample size can only be derived from accurate prior information [...] The sample size is therefore a planning parameter whose quality depends on the accuracy of the prior information. One cannot expect more from a planning procedure than the best possible use of prior information.” (Bock, 1998)

Motivation
(full factorial design)

- Goal: investigate treatment X with >1 other fixed factors (sex, dosage,…)
- Problem:
  - importance of the other fixed factors (interaction effects) for statistical analysis often not clear to the applicant of the animal experiment
  - applicable prior data for sample size calculation
- in practice of translational animal experiments: 2 means and 1 SD from transferable treatments (factor A with 2 levels)
  → Missing but needed: information about other fixed factors within the applicable data

→ solution proposal: simulations of possible factorial designs to evaluate a valid sample size

“Standard” sample size calculation

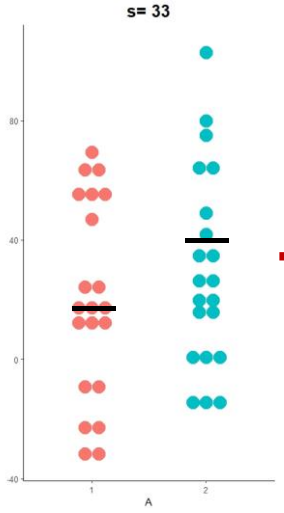
- Example: final design: 2x3 full factorial design (6 groups)
  information: $\mu_1 = 20$, $\mu_1 = 40$, $\sigma = 20$

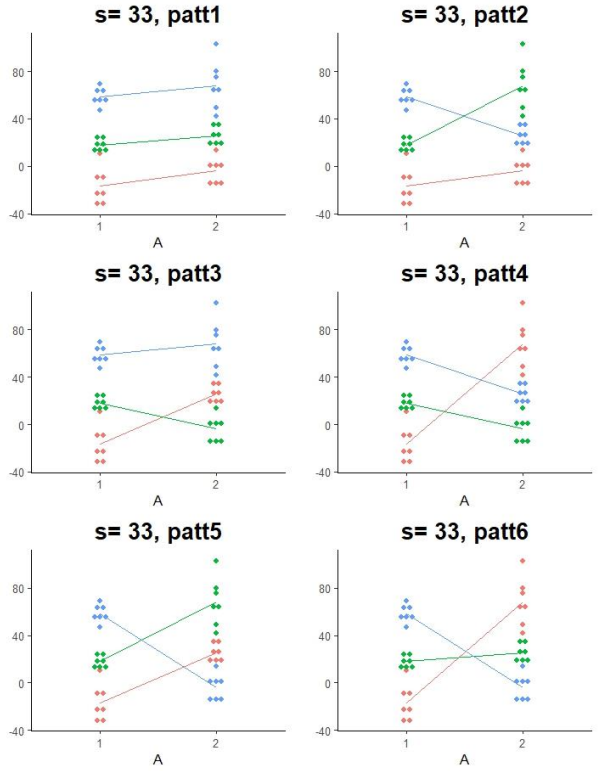
- $\alpha = 0.05$ (two-sided), $1 - \beta = 0.8$
- Resource equation method (Mead, 1990): $df_\epsilon = 10 - 20$ $[= 2 * 3(n-1)]$ → $N = 12$ $(n_{2x3} = 3)$ or $N = 18$ $(n_{2x3} = 4)$
- Rule of thumb: at least 3 animals per combination of the factor levels (Bate and Clark, 2014) → $N = 18$
- Unpaired t-test: $\mathrm{N_t} = 34$ $(n_t = 17^*)$ → $N = 36$ $(n_{2x3} = 6)$
- Wilcoxon-Mann-Whitney: $N_W = 40$ $(n_W = 20^*)$ → $N = 42$ $(n_{2x3} = 7)$

* by nQuery Advanced 9.3.1

Full factorial design – simulations

- different distribution assumptions of the simulated data (normal, lognormal, laplace, exponential)
- different scaling factors s (variation) and mean differences of the simulated data
- different number of other fixed factors (B,*C*) [besides factor A with 2 levels]
- with different number of levels (2,3,4,*5*)
- and different interaction patterns!





naïve arrangement
of additional factor levels

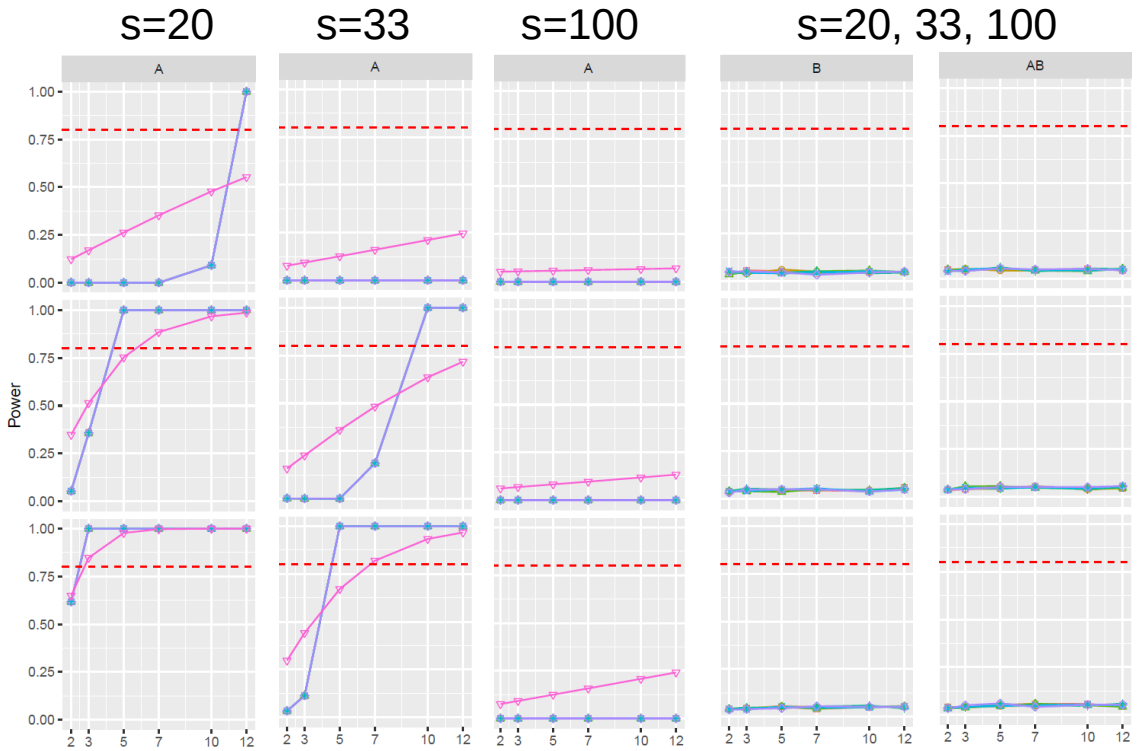
Full factorial design – results (1/4)

random
~normal

ΔMean=10

ΔMean=20

ΔMean=30



Power 80%

simulated observations per cell

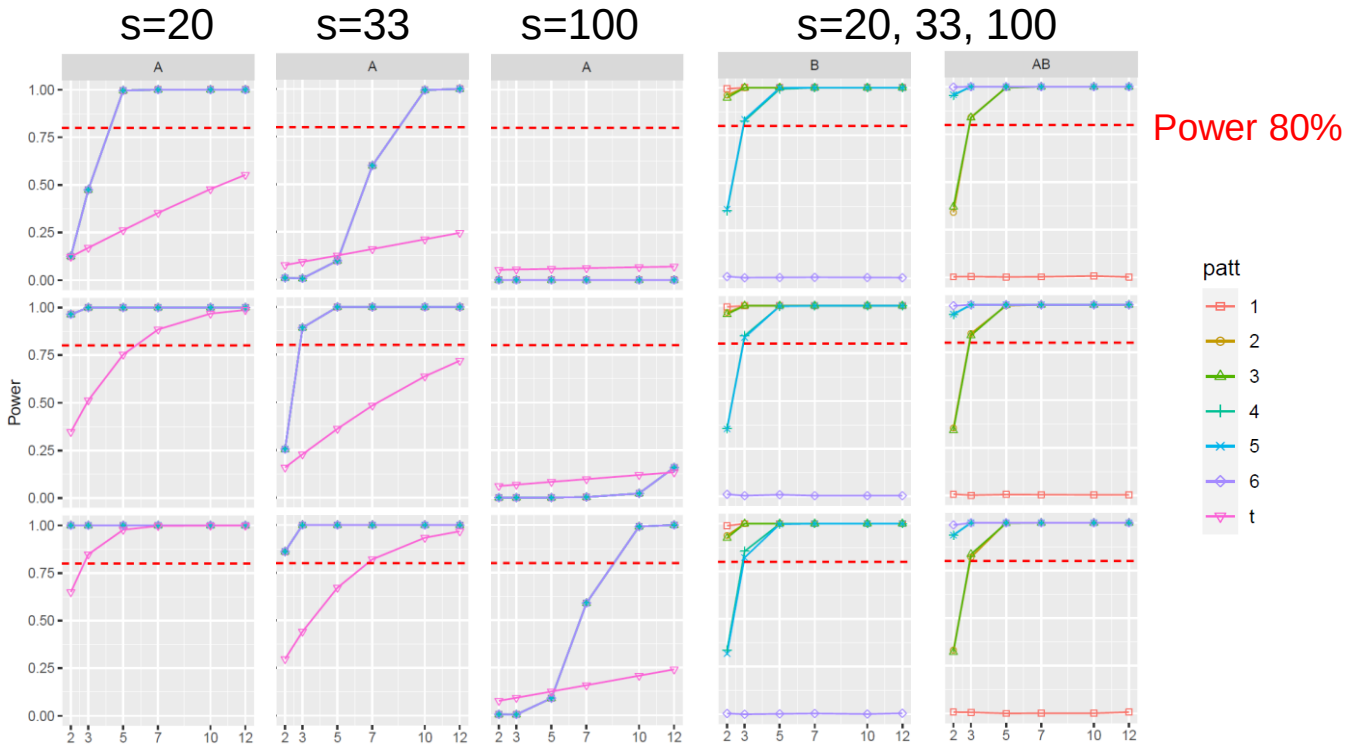
Full factorial design – results (2/4)

naïve
~normal

ΔMean=10

ΔMean=20

ΔMean=30



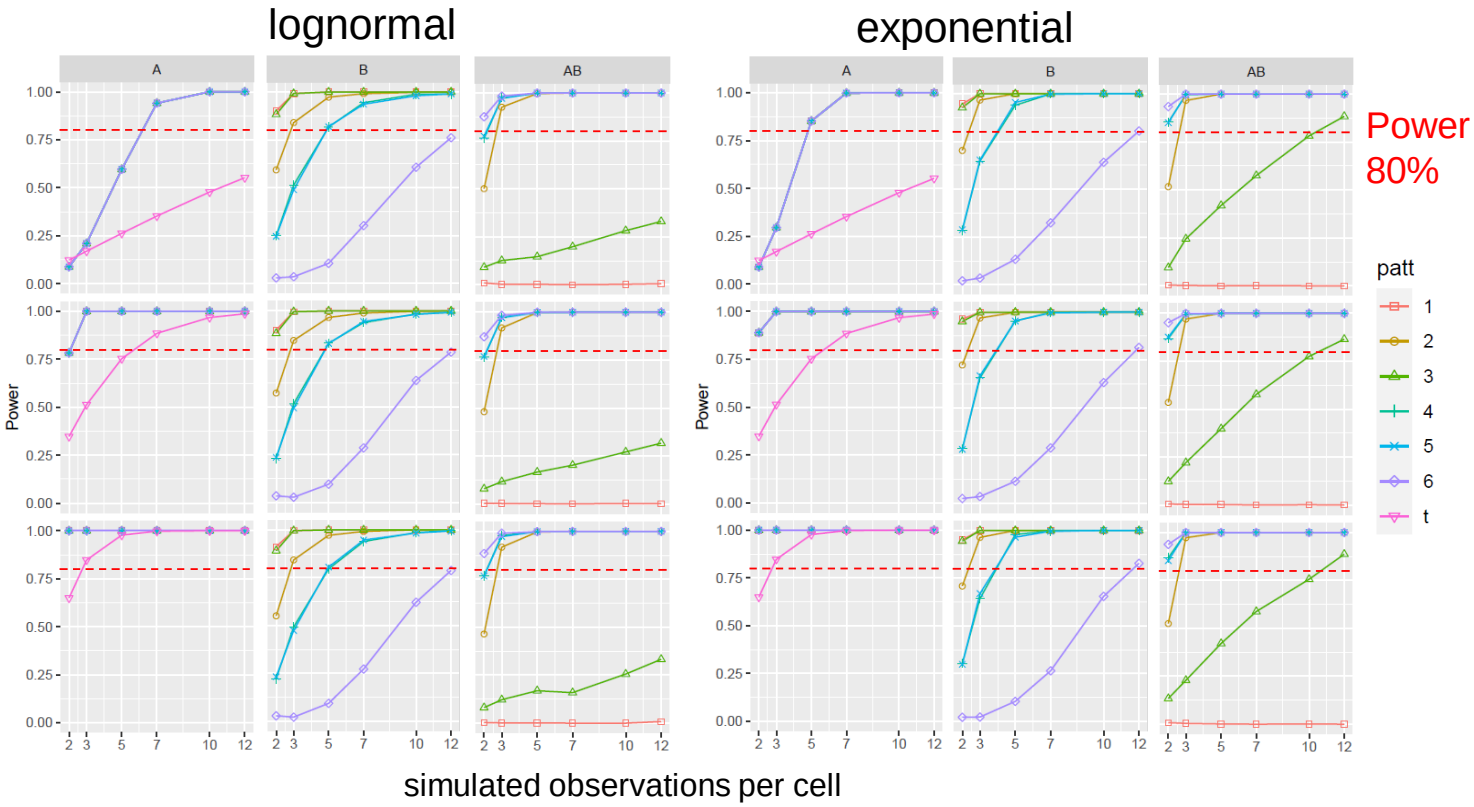
Full factorial design – results (3/4)

naïve, s=20
~non-normal

ΔMean=10

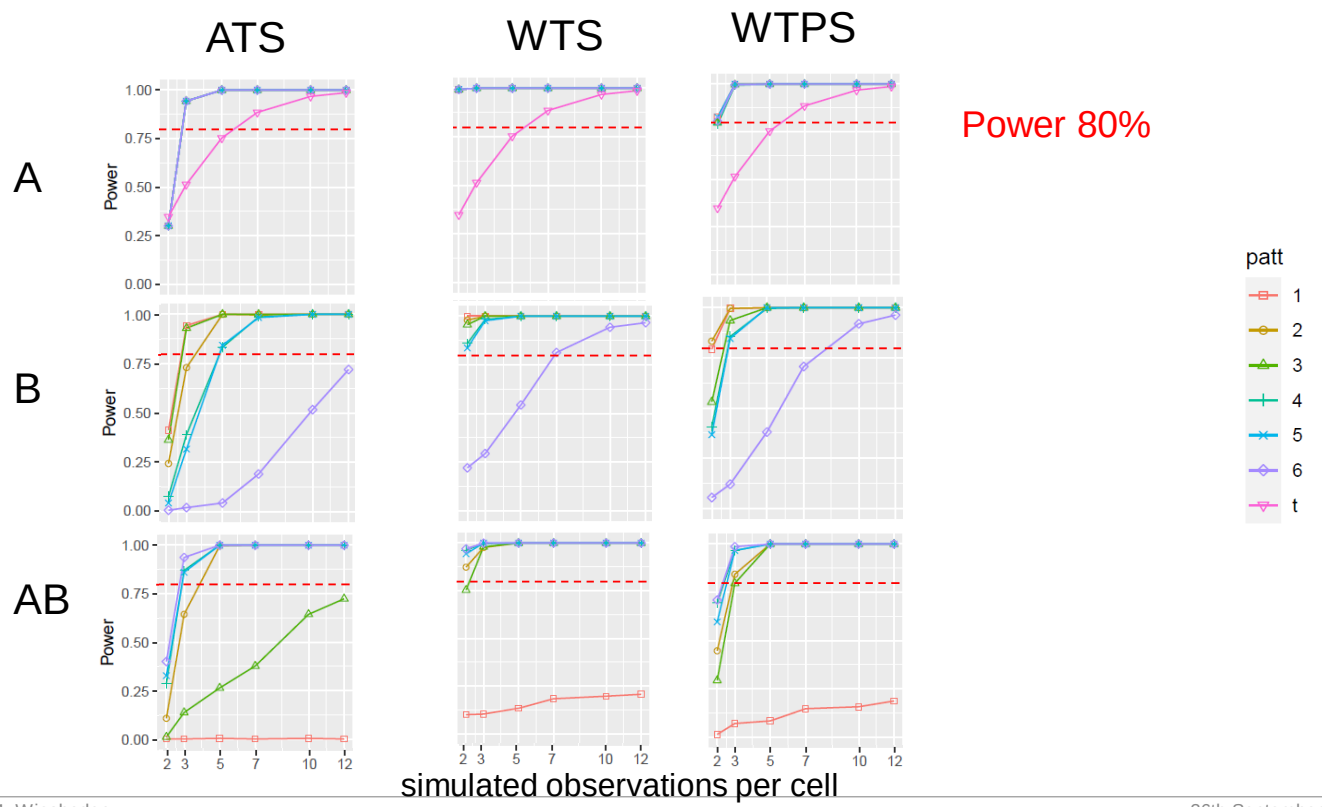
ΔMean=20

ΔMean=30



Full factorial design – results (4/4)

naïve, s=20,
ΔMean=20
~exponential



Power 80%

Summary, discussion, limitations

- simulations of possible interaction patterns within the given/assumed data can help to evaluate a meaningful sample size
  - help for applicant of the animal experiment
  - explorative vs. confirmatory analysis

- check:
  - full factorial design assumptions after naïve arrangement of additional factor levels
  - simulated data in the range of the outcome

- increasing complexity with more factors and levels

References

- Bate, S. T., and Clark, R. A. *The design and statistical analysis of animal experiments* (Cambridge university press, 2014).
- Bock, J. *Bestimmung des Stichprobenumfangs für biologische Experimente und kontrollierte klinische Studien* (Oldenbourg, Munich; Vienna, 1998)
- Cohen, J. *Statistical Power Analysis for the Behavioral Science* (Lawrence Erlbaum Associates, 1988).
- Dean, A., Voss, D. and Draguljic, D. *Design and Analysis of Experiments* 2nd ed. (Springer, New York, 2017).
- Mayer, B. and Muche, R. Die limitierte Aussagekraft formaler Fallzahlplanung im Rahmen von Tierversuchen der medizinischen Grundlagenforschung. In: *Tierärztliche Praxis Ausgabe K: Kleintiere/Heimtiere* 41.06. 367–374 (2013).
- Mead, R. *The design of experiments: statistical principles for practical applications* (Cambridge university press, 1990)
- Shaw, R., Festing, M. F. W., Peers, I. and Furlong, L.. Use of factorial designs to optimize animal experiments and reduce animal use. In: *ILAR journal* 43, 223-232 (2002).

Thanks for your attention

Dr. Andreas Allgöwer

andreas.allgoewer@uni-ulm.de

+49(0)731/50-26911