

THESIS PROPOSAL (REVISED)

**Interim sample size re-estimation for adequately powered
series of N -of-1 trials**

Daphne Weemering (3239480)

Supervisor: Peter van de Ven

Methodology and Statistics for the Behavioral, Biomedical and Social Sciences

Utrecht University

November 9, 2021

Word count: 750

Journal of publication: Statistics in Medicine

1 Introduction

1.1 The problem

Randomized controlled trials (RCTs) often require a relatively large sample size to have adequate power for detecting an effect in the population. This is not feasible for small populations, such as populations of people with a rare disease.

A methodological framework that can offer a solution in this setting, is the N -of-1 trial. A N -of-1 trial is a within-subject randomized controlled multiple crossover trial, where a patient repeatedly receives the experimental and control intervention in a random order. By combining several N -of-1 trials, we are able to estimate the population treatment effect.

The number of subjects and cycles within a N -of-1 trial is determined beforehand by means of sample size calculations. These calculations are challenging because important parameters such as the various variances are generally unknown before the start of the studies. If unrealistic assumptions are made with regard to these parameters, the study may be over- or underpowered.

1.2 What is and what isn't known?

Sample size formulas for N -of-1 trials were derived by Senn (2019) for both random- and fixed-effects models. For this thesis the interest lies with random-effects models, since the main objective of aggregating N -of-1 trials is to make inferences to the population. Random-effects models do require a larger sample size (Senn, 2019).

To conquer the problem of unrealistic assumptions for unknown parameters in sample size calculations, interim sample size re-estimation can be considered. Interim sample size re-estimation uses data that is already obtained to estimate the unknown parameters, and uses these estimates to recalculate the sample size. Proschan (2009) established sample size re-estimation methods for standard RCTs based on estimates of the treatment effect and on estimates of nuisance parameters. For this thesis, the latter approach will be applied. Wych et al. (2019) investigate the application of sample size re-estimation in crossover trials, a methodology similar to N -of-1 trials. They acknowledge that not much is known about sample size re-estimation for trials other than more 'standard' trials.

The application of interim sample size re-estimation in N -of-1 trials has not yet been investigated, and no specific methods and guidelines have been established. Also, it is yet unclear what the minimally required sample size is to carry out sample size re-estimation in N -of-1 trials.

1.3 Research question

The following two research questions will be addressed:

- How many patients and observations within patients must be observed in order to reliably (in terms of power) perform a sample size re-estimation for N -of-1 trials?

- How do series of N -of-1 trials with sample size re-estimation compare to series of N -of-1 trials without sample size re-estimation in terms of power and sample size?

When unknown variances are higher (/lower) compared to what is presumed beforehand, it is expected that a greater (/smaller) sample size is required. Therefore, we can expect that the approach applying sample size re-estimation will come closer to the desired power than the approach using a fixed sample size.

2 This project's strategy

To answer the research questions, the following steps should be taken:

- Look for existing sample size re-estimation methods and make these suitable for the use in N -of-1 trials.
- Evaluate these re-estimation methods in a simulation study where sample size re-estimation methods will be compared with N -of-1 trials with a fixed sample size. Methods will not be applied to empirical data.
 - The two approaches will be considered both in the situation where a priori assumptions about unknown parameters in the sample size calculations are realistic, and where these assumptions are not realistic.
 - Statistical power, i.e., fraction of iterations where the treatment effect is significant, and the (expected) sample size will be used for evaluating the performance of both approaches.

All analyses will be performed in R version 4.1.1 (R Core Team, 2021). A linear mixed model will be used to simulate data for a series of N -of-1 trials, with population model (following Senn (2019)):

$$\begin{aligned} d_{ij} &= \tau_i + \epsilon_{ij}, \quad i = 1, \dots, n, j = 1, \dots, k \\ \tau_i &\sim N(T, \psi^2) \\ \epsilon_{ij} &\sim N(0, 2\sigma^2) \end{aligned}$$

where i indexes the patient and j the cycle. ϵ_{ij} are random within-patient within-cycle disturbance terms. $\tau_1 \dots, \tau_n$ are the random treatment effects with mean T and variance ψ^2 .

Linear mixed models will be fitted using lme4 (Bates et al., 2021). The unknown parameters in this model are the variance of the treatment effect (ψ^2), the within-patient within-cycle variance (σ^2) and the average treatment effect (T). See table 1 for an overview of the parameters in the simulation.

The initial sample size (n) will be calculated as the sample size required to achieve 80% power under the hypothesized parameter values ψ^2 , σ^2 , the clinically relevant treatment effect (Δ) and a two-sided significance level of 5%. Sample size will be re-evaluated after fn subjects using the observed estimates for ψ^2 and σ^2 .

Table 1: Parameters in the simulation study

Description	Constant parameters	Varied parameters
Design parameters		
Fraction of initial sample size		$f = 0.25, 0.5, 0.75$
Cycles per patient	$k = 3$	
Power	$1 - \beta = 0.8$	
Two-sided significance level	$\alpha = 0.05$	
Clinically relevant difference	$\Delta = 1$	
Model parameters		
Within-patient within-cycle variance		$\sigma^2 = 0.25, 0.5, 1$
Variance of treatment effect		$\psi^2 = 0.5, 1, 2$
Average treatment effect	$T = \Delta = 1$	
Simulation parameter		
Number of simulations	$N = 10000$	
Outcome parameters		
Statistical power	Compare power of trial with sample size re-estimation in simulation to target of 80%	
Sample size	Compare sample size under re-estimation with sample size under true parameter values	

References

- Araujo, A., Julious, S., & Senn, S. (2016). Understanding variation in sets of N -of-1 trials. *PLoS ONE*, 11(12), e0167167. <https://doi.org/10.1371/journal.pone.0167167>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2021, June 22). **Package ‘lme4’** (Version 1.1-27.1). <https://cran.r-project.org/web/packages/lme4/lme4.pdf>
- Chen, X., & Chen, P. (2014). A comparison of four methods for the analysis of N -of-1 trials. *PLoS ONE*, 9(2), e87752. <https://doi.org/10.1371/journal.pone.0087752>
- Duan, N., Kravitz, R., & Schmid, C. (2013). Single patient (N -of-1) trials: A pragmatic clinical decision methodology for patient-centered comparative effectiveness research. *Journal of Clinical Epidemiology*, 66(1), S21–S28. <https://doi.org/10.1016/j.jclinepi.2013.04.006>
- Golkowski. (2014). Blinded sample size re-estimation in crossover bioequivalence trials. *Pharmaceutical Statistics*, 13(1), 157–162. <https://doi.org/10.1002/pst.1617>
- Grayling, M., Mander, A., & Wason, J. (2018). Blinded and unblinded sample size reestimation in crossover trials balanced for period. *Biometrical Journal*, 60(1), 917–933. <https://doi.org/10.1002/bimj.201700092>

- Heyvaert, M., Moeyaert, M., Verkempynck, P., Van den Noortgate, W., Vervloet, M., Ugille, M., & Onghena, P. (2016). Testing the intervention effect in single-case experiments: A monte carlo simulation study. *The Journal of Experimental Education*, 85(2), 175–196. <https://doi.org/10.1080/00220973.2015.1123667>
- Lenth, R. (2001). Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3), 187–194. <https://doi.org/10.1198/000313001317098149>
- Müller, A., Brands, M., Van de Ven, P., Roes, K., Cornel, M., Karnebeek, C., Wijburg, F., Daams, J., & ... (2021). Systematic review of N -of-1 studies in rare genetic neurodevelopmental disorders: The power of 1. *Neurology*, 96(11), 529–540. <https://doi.org/10.1212/WNL.00000000000011597>
- Proschan, M. (2005). Two-stage sample size re-estimation based on a nuisance parameter: A review. *Journal of Biopharmaceutical Statistics*, 15(1), 559–574. <https://doi.org/10.1081/BIP-200062852>
- Proschan, M. (2009). **Sample size re-estimation in clinical trials**. *Biometrical Journal*, 51(1), 348–357. <https://doi.org/10.1002/bimj.200800266>
- R Core Team. (2021). **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Senn, S. (2015a). Mastering variation: Variance components and personalised medicine. *Statistics in Medicine*, 35(1), 966–977. <https://doi.org/10.1002/sim.6739>
- Senn, S. (2015b). Various varying variances: The challenge of nuisance parameters to the practicing biostatistician. *Statistical Methods in Medical Research*, 24(4), 403–419. <https://doi.org/10.1177/0962280214520728>
- Senn, S. (2019). **Sample size considerations for n -of-1 trials**. *Statistical Methods in Medical Research*, 28(2), 372–383. <https://doi.org/10.1001/jama.2018.18020>
- Stunnenberg, B., Raaphorst, J., Groenewoud, H., Statland, J., Griggs, R., Woertman, W., Stegeman, D., Timmermans, J., & ... (2018). Effect of mexiletine on muscle stiffness in patients with nondystrophic myotonia evaluated using aggregated N -of-1 trials. *Journal of the American Medical Association*, 320(22), 2344–2353. <https://doi.org/10.1001/jama.2018.18020>
- Wood, F., Howard, J., Finegold, J., Nowbar, A., Thompson, D., Arnold, A., Rajkumar, C., & Connolly, S. (2020). N -of-1 trial of a statin, placebo, or no treatment to assess side effects. *The New England Journal of Medicine*, 383(22). <https://doi.org/10.1056/NEJMc2031173>
- Wych, J., Grayling, M., & Mander, A. (2019). **Sample size re-estimation in crossover trials: application to the AIM HY-INFORM study**. *Trials*, 20(665). <https://doi.org/10.1186/s13063-019-3724-6>
- Zucker, D., & Denne, J. (2002). Sample-size redetermination for repeated measures studies. *Biometrics*, 58(1), 548–559.
- Zucker, D., Ruthazer, R., & Schmid, C. (2010). Individual (N -of-1) trials can be combined to give population comparative treatment effect estimates: Methodologic considera-

tions. *Journal of Clinical Epidemiology*, 63(1), 1312–1323. <https://doi.org/10.1016/j.jclinepi.2010.04.020>

Zucker, D., Schmid, C., McIntosh, R., M.W. an Agostino, Selker, H., & Lau, J. (1997). Combining single patient (*N*-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clinical Epidemiology*, 50(4), 401–410.

* *Notice that the references with a boldface title are mentioned in the body of the text. The remaining references concern other relevant material that may be useful for the thesis.*