

COMMENTARY

Open Access



# Adapting the randomised controlled trial (RCT) for precision medicine: introducing the nested-precision RCT (npRCT)

Nils Kappelmann<sup>1,2\*</sup> , Bertram Müller-Myhsok<sup>1,3</sup> and Johannes Kopf-Beck<sup>4</sup>

## Abstract

Adaptations to the gold standard randomised controlled trial (RCT) have been introduced to decrease trial costs and avoid high sample sizes. To facilitate development of precision medicine algorithms that aim to optimise treatment allocation for individual patients, we propose a new RCT adaptation termed the nested-precision RCT (npRCT). The npRCT combines a *traditional* RCT (intervention A versus B) with a *precision* RCT (stratified versus randomised allocation to A or B). This combination allows online development of a precision algorithm, thus providing an integrated platform for algorithm development and its testing. Moreover, as both the *traditional* and the *precision* RCT include participants *randomised* to interventions of interest, data from these participants can be jointly analysed to determine the comparative effectiveness of intervention A versus B, thus increasing statistical power. We quantify savings of the npRCT compared to two independent RCTs by highlighting sample size requirements for different target effect sizes and by introducing an open-source power calculation app. We describe important practical considerations such as blinding issues and potential biases that need to be considered when designing an npRCT. We also highlight limitations and research contexts that are less suited for an npRCT. In conclusion, we introduce the npRCT as a novel precision medicine trial design strategy which may provide one opportunity to efficiently combine traditional and precision RCTs.

**Keywords:** Randomised controlled trial, Nested-precision randomised controlled trial, Precision medicine

## Background

In the wake of precision medicine, adaptations of the “gold standard” randomised controlled trial (RCT) have been introduced to accelerate the identification of personalised treatments, decrease trial costs, and avoid unnecessarily high sample sizes. These include adaptation studies such as the platform, linked Bayesian adaptive randomisation, or leapfrog design that utilise early stopping rules or (biomarker-dependent) mid-trial changes

in allocation ratio [1–4]. We propose a new RCT adaptation that could facilitate the move towards precision medicine, which aspires to give personalised treatment suggestions based on participant characteristics or pathological markers.

## The problem of retrospective precision algorithm development

To move towards precision medicine, researchers need to identify participant characteristics or pathological markers that are prescriptive, meaning they are informative on whether any individual participant benefits more from one treatment over another. As prescriptive markers for different interventions are often not known, current research efforts often revolve around re-

\* Correspondence: [n.kappelmann@gmail.com](mailto:n.kappelmann@gmail.com); [nils\\_kappelmann@psych.mpg.de](mailto:nils_kappelmann@psych.mpg.de)

<sup>1</sup>Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Kraepelinstraße 2-10, 80804 Munich, Germany

<sup>2</sup>International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

analysing existing RCT data to identify pre-treatment prescriptive variables and define precision algorithms.

A promising example for precision algorithm development in psychiatry is the Personalised Advantage Index (PAI) [5]. The PAI was developed based on a combination of prescriptive pre-treatment variables and provides a quantitative preference score on whether intervention A or intervention B is likely to be the optimal treatment for any individual participant. While approaches such as the PAI are fruitful for the development of new stratification rules, previous efforts have largely been restricted to algorithm development using existing RCT data. If precision algorithms were to find their way into clinical care, however, they would ideally require prospective testing similar to how interventions are evaluated more generally; that is, while interventions need to show empirical superiority over placebo or current gold standard treatments, precision algorithms should exhibit empirical superiority over random allocation to interventions.

To address the problematic predominance of retrospective precision algorithm development and lack of prospective evaluation, we propose what we term nested-precision RCT (npRCT), which allows for integrated precision algorithm development and evaluation.

## Introducing the npRCT

### Design and statistical power

As depicted in Fig. 1a, the npRCT combines a *traditional* RCT, comparing intervention A versus B, with a *precision* RCT, comparing stratification to intervention A or B versus randomisation to A or B.

This RCT combination has a twofold purpose. First and as noted above, combining these two RCT types can provide an integrated platform that allows for sequential development and testing of a precision algorithm. Specifically, researchers can analyse collected data during the *traditional* RCT to see if any strong prescriptive variables are present that could guide treatment choice. For example, identification of pre-treatment prescriptive variables can be conducted using the PAI approach, which looks for treatment-moderating variables using different regression-based and machine learning techniques [6]. Ideally, any of such online analyses of traditional RCT data should be conducted by an independent trial collaborator/statistician that is not involved in day-to-day recruitment or treatment procedures, so as to prevent potential researcher biases. If prescriptive variables are identified that could optimise intervention allocation decisions, researchers can move on to the *precision* RCT stage, in which the algorithm is prospectively evaluated against random intervention allocation.

The second purpose of the npRCT is to offer a design that can save trial costs as compared to running traditional and precision RCTs independently. As both the

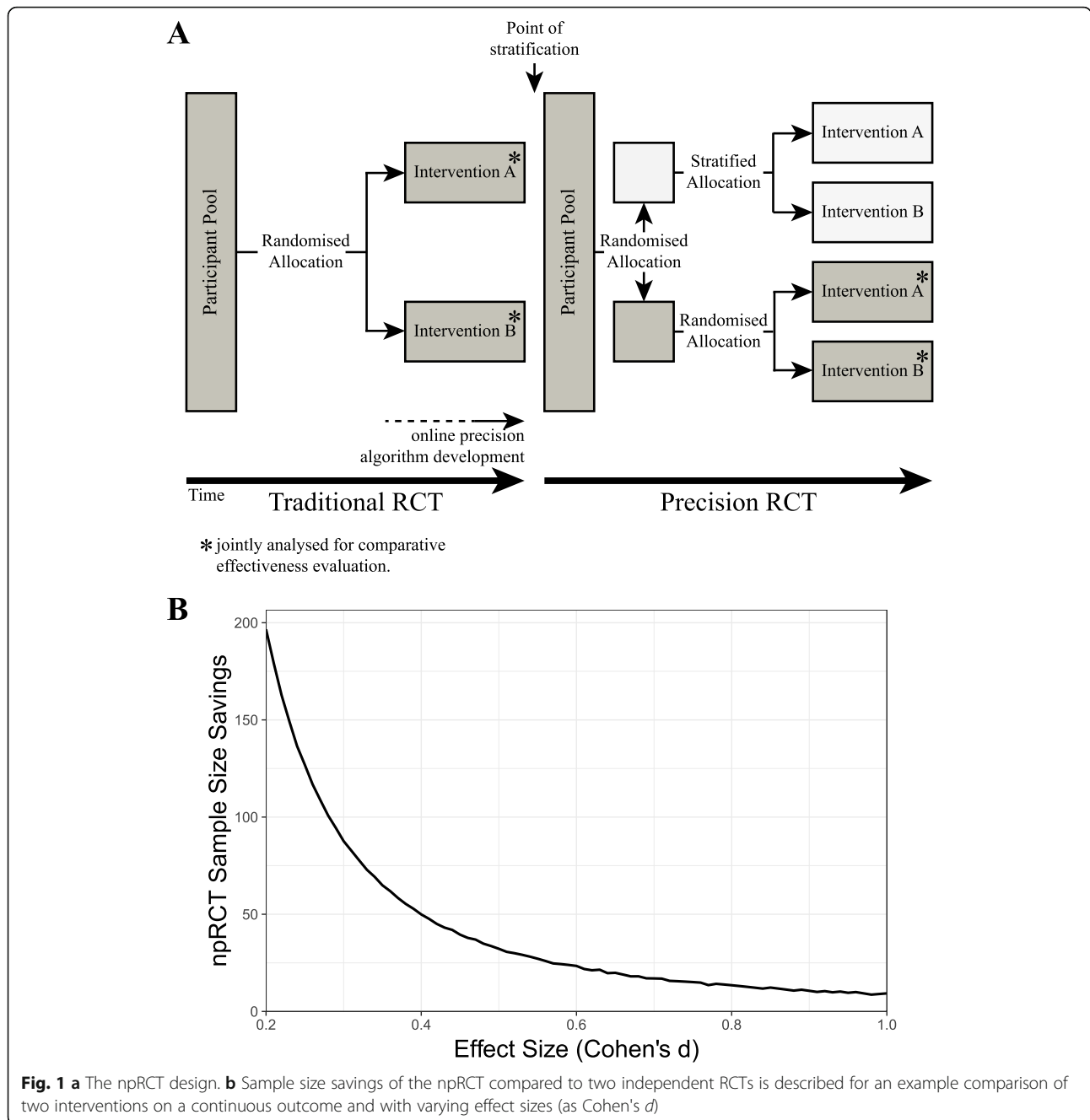
traditional and the precision RCT include participants randomised to interventions of interest, the traditional RCT can be understood as *nested* within the combined npRCT design. As such, all individuals randomised to interventions of interest can be jointly analysed to determine the comparative intervention effectiveness of interventions A and B. This nesting increases statistical power by reducing the sample size of the *traditional* RCT by 1/2 times the sample size of the *precision* RCT.

### Practical considerations

We want to highlight three major practical considerations that should be made when setting up an npRCT.

First, researchers should consider effect size and allocation ratio criteria a priori that determine if it is sensible to move on to the precision RCT stage or to maintain the traditional RCT. In terms of effect sizes, researchers need to specify minimum effect size benefits of the precision algorithm. If an effective algorithm can be identified based on this criterion, researchers can move on to the precision RCT stage. If an effective algorithm cannot be identified based on this criterion, researchers could maintain the traditional RCT design to focus on the comparative intervention effectiveness evaluation. Similarly, researchers should carefully consider how overall intervention main effects are weighed against treatment-moderating effects: Take a hypothetical example where intervention A is more effective than intervention B overall (i.e. a significant intervention main effect is present). Under this scenario, the algorithm may allocate a majority of participants to intervention A and only few participants to intervention B. If effectiveness differences are too large, a precision algorithm will lose its economic value as a precision medicine tool as it would make more sense to allocate all participants to intervention A irrespective of any prescriptive variables.

Second, researchers need to carefully consider potential blinding issues and biases arising from the npRCT, which are particularly likely if interventions cannot be concealed (e.g. through use of matched drug capsules). For example, researchers might learn how the precision algorithm allocates participants to interventions, which could lead to potential experimenter biases. Experimenter and expectation biases may also change from traditional to precision RCT stages. For example, participants might have a clear expectation that intervention A works best for them, so may assume (if allocated to intervention B) in the precision RCT stage that they were not assigned by the precision algorithm. To avoid expectation biases at the stratification stage, the aim should be that all individuals involved in the trial (i.e. participants, clinicians, and the research team) are blinded, so that knowledge about allocation group will not affect trial results.



Third, it will be important to set up a timeline that leaves sufficient time for online precision algorithm development. It is sensible to start precision algorithm development relatively late in the traditional RCT stage, so as to maximise sample sizes used for algorithm identification. However, the benefit of maximising sample sizes for algorithm identification should be weighed against practical time requirements to conduct analyses, especially if recruitment rate is high, to avoid potential delays between traditional and precision RCTs at the point of stratification.

**Power calculation app**

We have developed a free, open-source power calculation app for the npRCT that can be accessed via <https://nprct.shinyapps.io/nprct/>. This app includes power calculation strategies for (i) comparison of two groups on a continuous outcome (based on two-sample independent *t* tests), (ii) comparison of three or more groups on a continuous outcome (based on analysis of variance or ANOVA), and (iii) comparison of two or more groups on a binary or categorical outcome (based on  $\chi^2$  tests).

This power calculation app shows savings with this design (in terms of participants per group) compared to running the traditional and the precision RCT independently (see Fig. 1b). The power calculation app also provides the transition point from *traditional* to *precision* RCT (cf. point of stratification); that is, the point of recruitment at which one needs to change towards the precision RCT design. Considering a continuous outcome and comparison of two interventions, for frequently reported small, medium, and large effect sizes, the npRCT would require 196 (Cohen's  $d = 0.2$ ), 32 ( $d = 0.5$ ), and 13 ( $d = 0.8$ ) participants fewer *per group* compared to two independent RCTs (assuming  $\alpha = 0.05$  and power = 0.8). These statistics emphasise recruitment savings gained by combining/nesting RCTs under the npRCT design.

In the power calculation app, we provide explanations on further practical and ethical considerations. We also provide a walk-through research example for the npRCT. Source code of the npRCT app and power computation functions are freely available via github under <https://github.molgen.mpg.de/mpip/npRCT.app>, which we encourage researchers to use and adapt for more complex npRCT design setups (e.g. with traditional RCT allocation ratios deviating from 1:1 allocation).

### Advantages and limitations of the npRCT

We have proposed the npRCT as a strategy to integrate precision algorithm development and testing of stratified versus randomised allocation in a unified design and outlined potential advantages of this design in terms of participant savings. Despite these advantages, there are several limitations of the npRCT that researchers need to consider when deciding between clinical trial designs.

First, the npRCT does not allow answering of all research questions relevant for precision medicine. For instance, the precision RCT stage may enable investigators to obtain a “theoretical” estimate of precision algorithm effectiveness compared to random treatment allocation assuming that the biases mentioned under the [Practical considerations](#) section are minimised. In clinical practice, however, stratification algorithms may not be used as absolute decision rules to select between treatments. Instead, algorithmic decisions may be used as support decision-making tools to guide clinicians’ and patients’ choices, so patients and clinicians could opt to ignore the algorithmic recommendation as well. To obtain a “practical” estimate of precision algorithm effectiveness as a support decision-making tool would thus instead require an evaluation of randomised treatment allocation versus algorithmic treatment recommendation, which may not be a suitable research question for the npRCT due to unique blinding and bias considerations.

Second, the assumption of ethical favourability of the npRCT needs to be assessed individually for each research context. For instance, recent simulation work on precision algorithm development for major depression has suggested that precision algorithm development itself requires very large sample sizes for a range of plausible algorithmic effect sizes reported in the literature [7]. In this report, the authors have even suggested that algorithm development might realistically benefit from development in non-randomised study settings, only to be evaluated in a randomised trial subsequently. If npRCT power calculation proves unrealistic for a specific research context, we recommend the use of alternative trial designs including separate algorithm development and testing.

### Conclusion

We describe the npRCT, which combines two RCTs addressing research questions on (i) comparative intervention effectiveness (intervention A versus B) and (ii) precision treatment allocation (stratified allocation versus randomised allocation to A or B). Combining these two designs allows integrated precision algorithm development and testing. It also increases statistical power, which can be quantified using our free, open-source power computation app. Based on the resulting participant savings, the npRCT may be ethically and economically favourable to conducting two RCTs independently for certain research contexts. As such, we hope the npRCT may constitute one means of fostering prospective testing of precision medicine algorithms.

### Acknowledgements

Not applicable.

### Authors’ contributions

NK conceived and designed the npRCT and related power analysis, programmed the npRCT app, created the figure, and wrote the initial draft of the letter. JKB and BMM supervised this process and critically reviewed the letter. The authors read and approved the final manuscript.

### Funding

This project is funded by the Max Planck Institute of Psychiatry. NK is supported by the International Max Planck Research School for Translational Psychiatry.

### Availability of data and materials

As described in the manuscript, we provide a free, open-source power calculation app for the npRCT design under <https://npRCT.shinyapps.io/npRCT/>. Code for the app and npRCT power calculation functions can be retrieved from <https://github.molgen.mpg.de/mpip/npRCT.app>.

### Ethics approval and consent to participate

Not applicable (as no participant data have been collected).

### Consent for publication

Not applicable (as no participant data have been collected).

### Competing interests

NK, BMM, and JKB have nothing to disclose.

### Author details

<sup>1</sup>Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Kraepelinstraße 2-10, 80804 Munich, Germany. <sup>2</sup>International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany. <sup>3</sup>Department of Health Data Science, University of Liverpool, Liverpool, UK. <sup>4</sup>Max Planck Institute of Psychiatry, Munich, Germany.

Received: 21 April 2020 Accepted: 10 December 2020

Published online: 06 January 2021

### References

1. Garralda E, Dienstmann R, Piris-Giménez A, Braña I, Rodon J, Tabernero J. New clinical trial designs in the era of precision medicine. *Mol Oncol*. 2019; 13:549–57 Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6396357/pdf/MOL2-13-549.pdf>. Cited 2019 Mar 31.
2. Blackwell SE, Woud ML, Margraf J, Schönbrodt FD. Introducing the leapfrog design: a simple Bayesian adaptive rolling trial design for accelerated treatment development and optimization. *Clin Psychol Sci*. 2019;1–22 Available from: <http://journals.sagepub.com/doi/10.1177/2167702619858071>.
3. Hobbs BP, Chen N, Lee JJ. Controlled multi-arm platform design using predictive probability. *Stat Methods Med Res*. 2018;27:65–78 Available from: <http://journals.sagepub.com/doi/10.1177/0962280215620696>.
4. Wason JMS, Abraham JE, Baird RD, Gournaris I, Vallier A-L, Brenton JD, et al. A Bayesian adaptive design for biomarker trials with linked treatments. *Br J Cancer*. 2015;113:699–705 Nature Publishing Group. Available from: <http://www.nature.com/articles/bjc2015278>.
5. DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. The personalized advantage index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS One*. 2014;9:e83875 Cho WCS, editor. Available from: [www.plosone.org](http://www.plosone.org). Cited 2018 Jul 16.
6. Cohen ZD, Kim TT, Van HL, Dekker JJM, Driessen E. A demonstration of a multi-method variable selection approach for treatment selection: recommending cognitive-behavioral versus psychodynamic therapy for mild to moderate adult depression. *Psychother Res*. 2020;30:137–50. <https://doi.org/10.1080/10503307.2018.1563312> Taylor & Francis.
7. Luedtke A, Sadikova E, Kessler RC. Sample size requirements for multivariate models to predict between-patient differences in best treatments of major depressive disorder. *Clin Psychol Sci*. 2019;7:445–61 Available from: <http://journals.sagepub.com/doi/10.1177/2167702618815466>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

