JAMA Pediatrics | Special Communication

Experimental Designs to Optimize Treatments for Individuals
Personalized N-of-1 Trials

Karina W. Davidson, PhD, MASc; Michael Silverstein, MD, MPH; Ken Cheung, PhD; Rocco A. Paluch, MA; Leonard H. Epstein, PhD

Conventional randomized clinical trials (RCTs) compare treatment effectiveness to provide support for evidence-based treatments that can be generalized to the average patient. However, the information obtained from RCTs may not always be useful for selecting the best treatment for individual patients. This article presents a complementary approach to identifying optimized treatments using experimental designs that focus on individuals. Personalized, or N-of-1, designs provide both a comparative analysis of treatments and a functional analysis demonstrating that changes in patient symptoms are likely because of the treatment implemented. This approach contributes to the zeitgeist of personalized medicine and provides clinicians with a paradigm for investigating optimal treatments for rare diseases for which RCTs are not always feasible, identifying personally effective treatments for patients with comorbidities who have historically been excluded from most RCTs, handling clinical situations in which patients respond idiosyncratically (either positively or negatively) to treatment, and shortening the time lag between identification and implementation of an evidence-based treatment. These designs merge experimental analysis of behavior methods used for decades in psychology with new methodological and statistical advances to assess significance levels of changes in individual patients, and they can be generalized to larger populations for meta-analytic purposes. This article presents a case for why these models are needed, an overview of how to apply personalized designs for different types of clinical scenarios, and a brief discussion of challenges associated with interpretation and implementation of personalized designs. The goal is to empower pediatricians to take personalized trial designs into clinical practice to identify optimal treatments for their patients.

JAMA Pediatr. doi:10.1001/jamapediatrics.2020.5801
Published online February 15, 2021.

Author Affiliations: Center for Personalized Health, Northwell Health, New York, New York (Davidson); Donald and Barbara Zucker School of Medicine at Hofstra University/Northwell Health, Hempstead, New York (Davidson); Boston University, Boston, Massachusetts (Silverstein); Mailman School of Public Health, Columbia University, New York, New York (Cheung); Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, New York (Paluch, Epstein).

Corresponding Author: Karina W. Davidson, PhD, MASc, Center for Personalized Health, Northwell Health, 130 E 59th St, Ste 14C, New York, NY 10022 (kdavidson2@northwell.edu).

One of the tenets of personalized experimental medicine is that treatments that best benefit each patient should be identified. Parallel-group randomized clinical trials (RCTs) either provide evidence on how groups of patients respond to a treatment vs a control or compare alternative treatments, but they do not allow the identification of optimized treatments for an individual. An alternative experimental method and clinical decision approach is the personalized, or N-of-1, trial. Some consider such designs to be the pinnacle of the strength of evidence hierarchy (**Box**),[1] as they provide evidence for the assessment of net benefit (total benefits minus total harms) of treatments for individual patients. Educating clinicians and clinician-scientists about personalized trials provides a novel approach to improving pediatric care. Recent advances in remote, real-time passive monitoring, data analytics, and statistical approaches as well as easy access to placebo pills all warrant a resurgence of interest in these designs.[2,3] Unique benefits of these designs are their usefulness in determining optimal therapy for patients with rare diseases and improving care for patients with comorbidities who do not meet criteria for usual RCTs[4,5] or who experience unusual adverse effects or idiosyncratic treatment responses. The approach may also be useful in accelerating the transition from identification of evidence-based treatment to implementation of effective care. These designs can also test the effectiveness of combining different types of interventions (eg, behavioral and pharmaceutical treatments), as joint use of 2 treatments may need tailoring to be most effective for an individual patient.

Parallel-group randomized efficacy trials usually compare alternative treatments, or treatments with controls, and maximize the internal validity of findings across treatment groups. Parallel-group effectiveness trials are conducted with a wider range of participants, often in real-world settings, with the goal of bolstering external validity or extending the applicability of findings to broader populations. While different in their emphases, both types of RCT compare the average response to the treatment with comparator conditions—thereby accounting for variability in treatment response—with the goal of extrapolating optimal treatment for patients with characteristics similar to those of the study participants.

If there is minimal variability in response to treatment in an RCT, the best prediction of the magnitude of treatment benefit for an individual patient will be that estimated from the overall trial—the so-called main effect. However, there is almost always heterogeneity of treatment effects (HTE) or variability in the balance of benefits and harms of treatment found for different patients, both subgroups and individuals, included in the RCT.[6] HTE may include some patients with idiosyncratic treatment responses to initial treat-

ments, and these are difficult to ascertain from reports of parallel-group RCTs. In one empirical analysis, for example, only a minority of participants in RCTs experienced the average treatment benefit reported while most participants did not.[7] This type of finding from parallel-group RCTs is ubiquitous and suggests that there is substantial uncertainty about extrapolating net treatment benefit for individual patients.[8]

In addition to the frequency of HTE in clinical trials, there are several other reasons to consider alternative designs. First, RCTs often require hundreds of patients to have enough statistical power to precisely identify the treatment effect estimate and are thus infrequently conducted on patients with rare diseases. Second, many RCTs have strict entrance criteria, which a specific patient may not meet. This may be because of age, sex, race/ethnicity, or comorbidities and may force clinicians to extrapolate results from one population to another with little evidence to guide such decisions. Moreover, there is a wide temporal window between identification of an evidence-based treatment and adoption of that treatment in practice. RCTs often take an extraordinary amount time to be completed.[9] This means that discovery of potential benefits or harms is delayed for years before treatment is available. Personalized trials can provide tools for individual clinicians to test new treatments with patients in their care more rapidly and efficiently, and pediatricians may be more likely to adopt effective treatments if they can prove to themselves that these treatments work for their patients.

Introducing Personalized Trials

Personalized trials are single-participant, multiple-time-period, active-comparator crossover trials that can be randomized and masked.[10,11] As an alternative to extrapolating the results of a conventional trial to an individual, N-of-1 trials provide a clinician with an empirical answer about an optimal treatment for a specific patient. This type of personalized trial has been considered more rigorous than a systematic review of multiple RCTs for making evidence-based treatment decisions (Box), as systematic reviews still require clinicians to extrapolate results to their current patient.[12]

Personalized trials are part of a family of single-case designs that derive from the experimental analysis of behavior, and they have served as the platform for many current evidence-based treatments in psychology and medicine.[13-15] Single-case designs use experimental methods to maximize the likelihood that changes in outcome variables are caused by the intervention. These methods have been adapted into personalized trial designs, which are a specific form of randomized or balanced experiments characterized by periodic and preplanned switching from active treatment to placebo or between active treatments (withdrawal-reversal designs).[16] Whereas conventional RCTs randomize persons, personalized trials randomize time periods for an individual patient. This type of design addresses concerns with regard to averaging treatment effects across many patients in the presence of known or unknown HTE, external validity–extrapolation issues, the exclusion of patients with rare diseases or comorbidities, and the time lag between treatment discovery and implementation.

Box. Original Hierarchy of Strength of Evidence for Treatment Decisions[1]

- N-of-1 randomized trials
- Systematic review of randomized clinical trials
- Single randomized clinical trials
- Systematic review of observational studies addressing patient-important outcomes
- Single observational study addressing patient-important outcomes
- Physiologic studies
- Unsystematic clinical observations

General Principles of Personalized Designs

There are several common characteristics of personalized trial designs. Personalized designs trade collecting a few data points from many people for a detailed examination of the relationship between treatment and outcome for a single individual. Because of this, it is important to use outcome measures that can be evaluated frequently and show rapid change when a treatment is implemented or withdrawn. If a drug is studied, it should be administered under randomized, placebo-controlled conditions with appropriate return-to-baseline symptomatology or drug washout periods.

Personalized designs work well for patients who have chronic, stable, or slowly deteriorating conditions. These designs can also be used for prevention when there is a continuous risk that has not yet crossed the diagnostic threshold, such as abnormal blood pressure or hemoglobin A1c levels in the prediabetes range. Thus, these designs can be used to prevent the clinical manifestation of a disease in some circumstances. They can also be used to identify an optimized treatment plan for a patient as they begin a treatment regimen or when a change in an established but ineffective treatment is being considered. However, given that these are experimental designs and that treatment comparators are delivered serially over randomly varying time periods, patients who require immediate treatment or have urgent medical needs may not be well suited for this type of experiment.

Uses for a Personalized Trial

One broad use of personalized trials is for patients with common conditions that do not yet have universally beneficial or evidence-based treatments. Consider obesity, pain, asthma, irritable bowel syndrome, or a variety of for a variety of difficult-to-treat behavioral and biomedical disorders—each of these is amenable to a personalized trial because there is clinical uncertainty as to the best treatment, conflicting evidence, known idiosyncratic responses to common treatments across individual patients, and clinically important end points that are easily measurable over a predictable time frame.

A second broad use of personalized trials is for rare diseases for which there are insufficient numbers for an RCT to provide stable findings for evidence-based solutions and for which effective treatments are not known. Data from personalized trials can often be

combined to provide an evidence base for treatments for children with rare diseases.

Third, personalized trials can be used when a child has multiple comorbidities and consequently is taking a polypharmaceutical regimen with suspected iatrogenic effects. In this case, the family and the clinician may want to investigate the safety of removing suspected medications from the regimen. Children with multiple morbidities are not often recruited for RCTs, and results of evidence-based treatments may not generalize to patients with multiple comorbidities. Additionally, concerns about the efficacy of brand-name vs generic medications and new adaptation of conventional treatments can be tested with personalized trials. While the goal of most experiments is to test whether treatments differ in their outcomes, there are instances in which a noninferiority comparison would be useful, such as when evaluating a generic formulation or lower dose of a medication.

Fourth, personalized trials can serve as proof-of-concept studies to test new interventions that speed up development of innovative treatments to test in RCTs. In this way, there can be beneficial feedback loops between personalized trials designed for discovery and the conduct of parallel-group RCTs to ensure a generalizable treatment benefit. Personalized trials are thus recognized as important steps in translating basic science into new clinical interventions.[17]

Statistical Considerations for Analyzing a Personalized Trial

A basic step in analyzing data from a personalized trial is to examine the results graphically. This first step is useful for understanding the relationship between treatment and outcome or the relative efficacy of various treatments. For some personalized trials, graphical analysis may provide such obvious results as to sufficiently inform conclusions about optimal treatment. In most instances, however, statistical analyses are needed to compare the treatments. Time-series analyses can be used to leverage the frequent outcome assessments in a personalized trial.[18] For example, autoregression models evaluate treatment effects while accounting for the association between successive outcomes and may be used to estimate the extent of carryover effects.[19] Advanced time-series methods have also been developed for trials with high-volume data measured via wearable devices[20]—such as actigraphs, step counters, and heart rate monitors—as they become available. These analytic approaches provide valid treatment comparisons and contrast with comparative analyses in RCTs where the focus is on estimating the average treatment effect based on few data points collected from each individual.

Similar to parallel-group RCTs, statistical issues, such as power and multiplicity adjustments, need to be addressed when planning a personalized trial. The power to detect a treatment effect specific to an individual depends on the association among outcomes as well as the magnitude of treatment effect. Generally, a larger number of measurements in a time series is required with higher associations across outcomes. In trials involving more than 2 treatments, it is necessary to adjust for multiple comparisons. Using a gate-keeping test can safeguard against false-positive findings,[21] while traditional methods, such as Bonferroni adjustment, are known to be overly conservative and to unnecessarily reduce power.

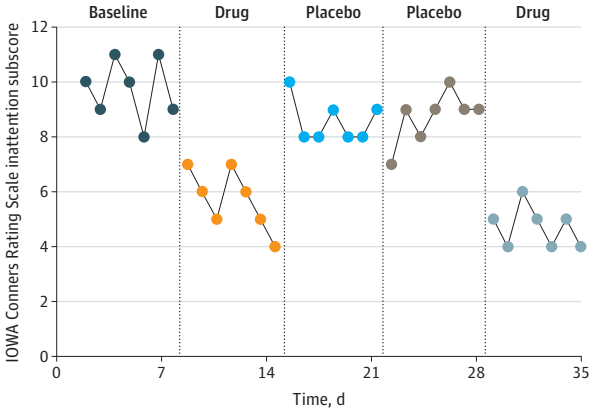
Finally, while a personalized trial focuses on developing optimized treatment for an individual patient, data can be combined across multiple patients to generate meta-analyses of personalized trials—just like conventional meta-analyses.[13,15,22] Specifically, results from individual patients can be combined in random-effects models and have the advantage of incorporating variability from different sources, including the overall treatment effect and patient-by-treatment interactions.[23] In addition, when combining several personalized trial series that compare treatments, it is possible to eliminate time trend when treatments are introduced in a randomized sequence in different trials.

Cases for Personalized Trials

We illustrate the basics of personalized trial designs using 3 cases of attention-deficit/hyperactivity disorder (ADHD)—2 hypothetical and 1 actual. In the first hypothetical case, the pediatrician is determining if pharmacological treatment is useful for a 12-year-old child newly diagnosed with ADHD based on daily parent ratings on the subscale of inattention from the Inattention and Overactivity With Aggression (IOWA) Conners Rating Scale (ICRS subscale; range, 1-20 points).[24] To ensure that any improvements in ADHD symptoms are due to the pharmaceutical treatment, the pediatrician includes a baseline that precedes the experimental period, ideally randomizing experimental periods and masking parent and child to the type of pharmaceutical intervention (drug or placebo), thus creating a multiperiod crossover design. The baseline is critical because it serves as a control for the trial. If testing a drug or treatment that has a long half-life, unlike those typically used to manage ADHD, then additional washout periods may be needed to ensure a return to baseline symptom level in the intervals between periods. A graphical presentation of daily ICRS subscales of inattention in an ABCCB trial is shown in **Figure 1**. With relatively high inattentive scores during baseline and placebo conditions, it is clear that the scores are consistently improved (ie, lower) during the drug period. This is an example where visual inspection may be sufficient to determine that pharmaceutical treatment has benefits.

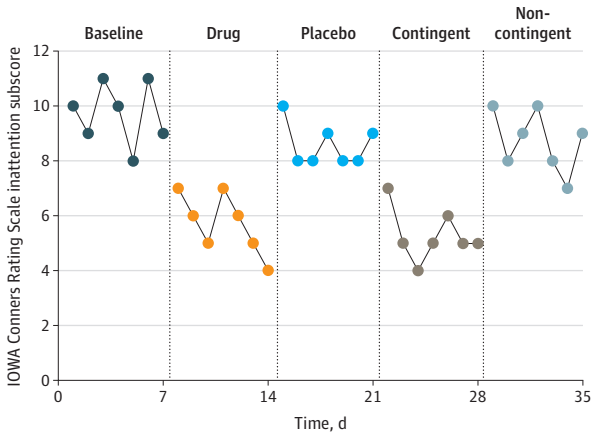
In the second hypothetical case, a 9-year-old patient was recently diagnosed with ADHD following a history of school problems. The family has concerns about long-term use of pharmaceutical treatment and wants to determine if a behavioral intervention might be effective. To ensure that the behavioral intervention works as well as or better than medication, the pediatrician chooses a similar design as in the first example, including a baseline and 4 experimental periods (ie, drug, drug placebo, contingent reinforcement intervention, and noncontingent reinforcement), and ideally randomizes said periods. In this case, masking or blinding to the condition is not possible for the behavioral intervention, but it is possible for the drug vs placebo periods. As treatments should ideally be reversible, the pediatrician chooses a contingent reinforcement behavioral intervention and a noncontingent control. During the contingent reinforcement condition, when the child achieves a score of 6 or lower on the ICRS in a 4-hour block, they are provided with a skipping rope and 20 minutes of physical activity reinforcement. In the noncontingent reinforcement condition, the child is provided with the physical activity reinforcer at randomly chosen times, independent of their behavior. In other periods of the trial, the activ-

Figure 1. Hypothetical Personalized Trial Results for a 12-Year-Old With Attention-Deficit/Hyperactivity Disorder



This model presents a hypothetical case of a 12-year-old patient newly diagnosed with attention-deficit/hyperactivity disorder. Inattention symptoms are assessed by parental Inattention and Overactivity With Aggression (IOWA) Conners Rating Scale[24] subscores using an ABCCB design to determine drug efficacy.

Figure 2. Hypothetical Personalized Trial Results for a 9-Year-Old With Attention-Deficit/Hyperactivity Disorder



This model presents a hypothetical case of a 9-year-old patient recently diagnosed with attention-deficit/hyperactivity disorder. Inattention symptoms are assessed by parental Inattention and Overactivity With Aggression (IOWA) Conners Rating Scale[24] subscores using an ABCDE design to determine drug efficacy.

ity reinforcer is not available. A sticker reward system could be used as an alternative to physical activity reinforcers. **Figure 2** depicts this trial, wherein both drug and behavioral intervention clearly improve the outcome over the placebo. Fitting an autoregression model demonstrates significant effects with both drug and contingent reinforcement compared with placebo; reductions in ICRS score were 2.8 points for drug use and 3.2 points for behavioral intervention. Furthermore, the 95% CI for comparing the scores during behavioral intervention and drug periods is −1.4 to 0.5, suggesting equivalence of the 2 modalities and providing empirical evidence for the family's preference for contingency management for their child.

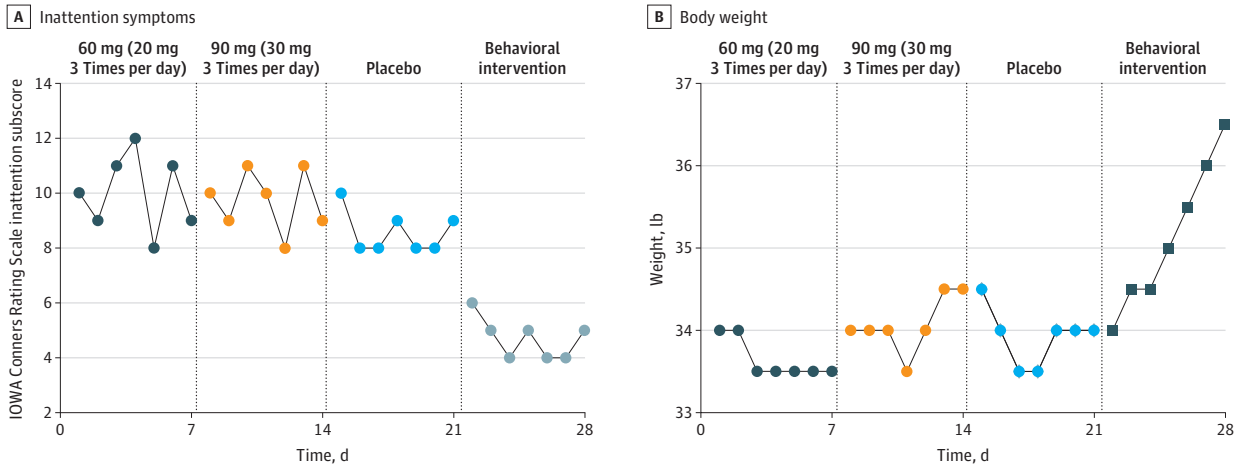
A few practical considerations should be noted in this case. First, because a behavioral program involves parent training for administration of a positively reinforcing contingency, parent education must be included in the overall treatment plan. Second, behavioral interventions often cannot be blinded, and it will be clear to the clinician and patient that a behavioral intervention is being implemented. Third, we have illustrated a relatively stable baseline, but in reality, symptoms may fluctuate prior to intervention, thereby requiring that the duration of the baseline be extended or reduced. Fourth, it would be possible to assess how best to integrate pharmacological and behavioral treatments if that is the goal of the pediatrician.

The third, actual case is a 7-year-old patient who weighed 34 lb and had been diagnosed with nonorganic failure to thrive and ADHD with both inattention and oppositional defiant components. The patient was taking 60 mg/d of methylphenidate, the maximum dosage approved by the US Food and Drug Administration. The patient's pediatrician, recognizing methylphenidate may have the off-target effect of appetite suppression, wanted to increase the methylphenidate dosage as an off-label use to determine if the child's inattentive behavior was responsive to increased doses. In the absence of a best course of treatment for this patient and a lack of RCT evidence because of comorbidities, a double-blind, randomized N-of-1 dose-finding trial was conducted, offering the flexibility to compare different modalities. **Figure 3** shows the patient's inattention scores and weights in a trial of current drug use, a higher-dose alternative, a placebo, and a behavioral contingency intervention. Because this child was already taking 60 mg/d of methylphenidate, that was the baseline condition. Visual inspection demonstrates behavioral intervention as the optimal treatment for both inattention improvement and weight gain. The drug was found to be significantly less effective than placebo or the behavioral intervention in terms of inattention scores and weight. Consequently, the drug was discontinued. Such a change in treatment for this child may not have been possible without the personalized trial, as the patient might have continued to take a drug that had an adverse effect on weight and little to no benefit for the attention symptoms. One of the strengths of personalized trials is that they can empirically inform the discontinuation of a current treatment that has shown no evidence of effectiveness.[19]

Implementation Considerations for Personalized Trials

When drugs are being studied, personalized trials require close coordination between a research-oriented pharmacist who can implement a randomized sequence of drug use, drug dosage variation, and placebo phases. Personalized trials also benefit from statistical analysis to determine whether differences in patient symptoms are significant, thus requiring close collaboration with a biostatistician. While the statistical analyses for individual patient data may require sophisticated statistical methods, there is an increased capability for a trained biostatistician to publish user-friendly web-based software for nonstandard analyses, such as the Shiny package (RStudio) for R versions 3.0.2 onward (The R Foundation). These statistical software or online statistical consultations that only require entry of an individual patient's data, often without the need for disclosing protected health information, can be developed to make analysis as easy as possible and

Figure 3. Personalized Trial Results for a 7-Year-Old With Attention-Deficit/Hyperactivity Disorder and Nonorganic Failure to Thrive



This model presents an actual case of a 7-year-old patient diagnosed with attention-deficit/hyperactivity disorder and nonorganic failure to thrive. Inattention symptoms are assessed by teacher Inattention and Overactivity With Aggression (IOWA) Conners Rating Scale[24] subscores (panel A) and body weight (panel B) using an ABCD design.

thereby lower the barriers to implementing personalized trials because of lack of access to a local biostatistical collaborator. Personalized trials also require collection of a lot of patient data as well as staff who can collect data, check data quality, and organize and present data. This challenge is not unique to personalized trials, but a greater use of transducers and reliable patient reporting methods to collect data on blood glucose level, blood pressure, activity, sleep, body weight, etc have been adapted for clinical work with the same requirements on staff.

In this type of trial, there are complex patient consent issues that may be different depending on the motivation of the trial. Treatment discovery is clearly considered research while optimal treatment identification may be considered best clinical practice. Questions of payment for engaging in a personalized trial may be appropriate in some cases and not in others. Informed consent is always needed, as is extended measurement.

Conclusions

Personalized trials using N-of-1 designs have been used for decades to identify effective interventions across disciplines, but they have not yet found a home in pediatrics. This approach is useful for rare diseases, for specialists who are working on developing innovative treatments for their patients, and for pediatricians in primary-care settings who want to ensure they are providing the best patient care given the current state of knowledge. It may also be useful for testing treatments in patients with comorbidities, suspected adverse effects from current treatments, or idiosyncratic treatment responses. Using these designs in clinical practice can empower clinicians to test established and new evidence-based treatments for patients in their practices, thus accelerating the time from treatment-benefit discovery to clinical implementation.

ARTICLE INFORMATION

Accepted for Publication: June 6, 2020.

Published Online: February 15, 2021. doi:10.1001/jamapediatrics.2020.5801

Author Contributions: Dr Cheung and Mr Paluch had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Davidson, Silverstein, Epstein.
Acquisition, analysis, or interpretation of data: Davidson, Cheung, Paluch.
Drafting of the manuscript: Davidson, Silverstein, Epstein.
Critical revision of the manuscript for important intellectual content: Davidson, Silverstein, Cheung, Paluch.
Statistical analysis: Cheung, Paluch, Epstein.
Obtained funding: Davidson.
Administrative, technical, or material support: Davidson.

Conflict of Interest Disclosures: Drs Davidson and Silverstein are members of the US Preventive Services Task Force. No other disclosures were reported.

Funding/Support: This work was supported by the National Library of Medicine of the National Institutes of Health grant R01LM012836; the Eunice Kennedy Shriver National Institute of Child Health and Human Development grants R01HD080292 and R01HD088131; the National Heart, Lung, and Blood Institute grant U01 HL131552; and the National Institute of Diabetes and Digestive and Kidney Diseases grant UH3 DK109543.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The views expressed in this article are those of the authors and do not represent the views of the National Institutes of Health, the US Department of Health and Human Services, or any other government entity. This article does not represent the views and policies of the US Preventive Services Task Force.

REFERENCES

1. Guyatt G, Haynes B, Jaeschke R, et al. Introduction: the philosophy of evidence-based medicine. In: Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-based Clinical Practice*. American Medical Association; 2002:3-11.

2. Davidson KW, Cheung YK. Envisioning a future for precision health psychology: innovative applied statistical approaches to N-of-1 studies. *Health Psychol Rev*. 2017;11(3):292-294. doi:10.1080/17437199.2017.1347514

3. Davidson KW, Peacock J, Kronish IM, Edmondson D. Personalizing behavioral interventions through single-patient (N-of-1) trials. *Soc Person Psychol Compass*. 2014;8(8):408-421. doi:10.1111/spc3.12121

4. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet*. 2012;380(9836):37-43. doi:10.1016/S0140-6736(12)60240-2

5. Stoll CRT, Izadi S, Fowler S, et al. Multimorbidity in randomized controlled trials of behavioral interventions: a systematic review. *Health Psychol*. 2019;38(9):831-839. doi:10.1037/hea0000726

6. Vijan S. Evaluating heterogeneity of treatment effects. *Biostat Epidemiol*. 2020;4(1):98-104. doi:10.1080/24709360.2020.1724003

7. Ioannidis JP, Lau J. The impact of high-risk patients on the results of clinical trials. *J Clin Epidemiol*. 1997;50(10):1089-1098. doi:10.1016/S0895-4356(97)00149-2

8. Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials*. 2010;11:85. doi:10.1186/1745-6215-11-85

9. Lauer MS, D'Agostino RB Sr. The randomized registry trial—the next disruptive technology in clinical research? *N Engl J Med*. 2013;369(17):1579-1581. doi:10.1056/NEJMp1310102

10. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy—randomized trials in individual patients. *N Engl J Med*. 1986;314(14):889-892. doi:10.1056/NEJM198604033141406

11. McLeod RS, Taylor DW, Cohen Z, Cullen JB. Single-patient randomised clinical trial. use in determining optimum treatment for patient with inflammation of Kock continent ileostomy reservoir. *Lancet*. 1986;1(8483):726-728. doi:10.1016/S0140-6736(86)91112-8

12. Davidson KW, McGinn T, Wang YC, Cheung YK. Expanding the role of N-of-1 trials in the precision medicine era: action priorities and practical considerations. *NAM Perspectives*. Published online December 10, 2018. doi:10.31478/201812d

13. Shaffer JA, Kronish IM, Falzon L, Cheung YK, Davidson KW. N-of-1 randomized intervention trials in health psychology: a systematic review and methodology critique. *Ann Behav Med*. 2018;52(9):731-742. doi:10.1093/abm/kax026

14. Kronish I, Falzon L, Hampsey M, Konrad B, Davidson K. Are N-of-1 trials useful for achieving therapeutic precision in the treatment of depression? a systematic review of N-of-1 trials for depressive symptoms. *Psychosom Med*. 2017;79(4):A159.

15. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Med Care*. 2011;49(8):761-768. doi:10.1097/MLR.0b013e318215d90d

16. Kravitz RL, Duan N, eds; The DEcIDE Methods Center N-of-1 Guidance Panel. *Design and Implementation of N-of-1 Trials: A User's Guide*. Agency for Healthcare Research and Quality; 2014.

17. Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What ever happened to N-of-1 trials? insiders' perspectives and a look to the future. *Milbank Q*. 2008;86(4):533-555. doi:10.1111/j.1468-0009.2008.00533.x

18. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis: Forecasting and Control*. 4th ed. John Wiley & Sons, Inc.; 2008. doi:10.1002/9781118619193

19. Kronish IM, Cheung YK, Shimbo D, et al. Increasing the precision of hypertension treatment through personalized trials: a pilot study. *J Gen Intern Med*. 2019;34(6):839-845. doi:10.1007/s11606-019-04831-z

20. Cheung YK, Hsueh PS, Ensari I, Willey JZ, Diaz KM. Quantile coarsening analysis of high-volume wearable activity data in a longitudinal observational study. *Sensors (Basel)*. 2018;18(9):E3056. doi:10.3390/s18093056

21. Zhong X, Cheng B, Qian M, Cheung YK. A gate-keeping test for selecting adaptive interventions under general designs of sequential multiple assignment randomized trials. *Contemp Clin Trials*. 2019;85:105830. doi:10.1016/j.cct.2019.105830

22. Kronish IM, Hampsey M, Falzon L, Konrad B, Davidson KW. Personalized (N-of-1) trials for depression: a systematic review. *J Clin Psychopharmacol*. 2018;38(3):218-225. doi:10.1097/JCP.0000000000000864

23. Jull A, Bennett D. Do N-of-1 trials really tailor treatment? *Lancet*. 2005;365(9476):1992-1994. doi:10.1016/S0140-6736(05)66678-0

24. Pelham WE Jr, Milich R, Murphy D, Murphy HA. Normative data on the IOWA Conners Teacher Rating Scale. *J Clin Child Psychol*. 1989;18(3):259-262. doi:10.1207/s15374424jccp1803_9