



Published in final edited form as:

J Clin Psychopharmacol. 2018 June ; 38(3): 218–225. doi:10.1097/JCP.0000000000000864.

Personalized (N-of-1) Trials for Depression: A Systematic Review

Ian M. Kronish, MD, MPH¹, Meghan Hampsey, BA¹, Louise Falzon, PGDipln¹, Beatrice Konrad, BS¹, and Karina W. Davidson, PhD¹

¹Center for Behavioral Cardiovascular Health, Columbia University Medical Center, 622 W. 168th St. New York, NY 10032

Abstract

Purpose/Background—Personalized (N-of-1) trials are single-patient, crossover-design trials that may be useful for personalizing the selection of depression treatments. We conducted a systematic review of published N-of-1 trials for depression to determine the feasibility and suitability of this methodology for personalizing depression care.

Methods/Procedures—Electronic databases were searched from database inception through October 2016. Studies were selected if they enrolled depressed patients, included a within-subject crossover design, and systematically assessed depressive symptoms during the N-of-1 trial.

Findings/Results—Five eligible studies reporting on 47 depressed patients (range 1 to 18 patients) were identified. Two studies were conducted among adults with treatment resistant depression, one among depressed inpatients, and two among patients from special populations (geriatric nursing home, HIV-encephalopathy). All studies evaluated the effects of pharmacologic treatments (methylphenidate, d-amphetamine, ketamine, and sulpiride). Three studies compared an off-label treatment with placebo, one study compared two off-label treatments, and one study compared escalating doses of an off-label treatment with placebo. All four studies with more than one participant demonstrated heterogeneous treatment effects. All studies produced data that could personalize treatment selection for individual patients. No studies reported on recruitment challenges, compliance with self-tracking, nor satisfaction with participation.

Implications/Conclusions—The feasibility of N-of-1 trials for depression was demonstrated for a limited number of second-line pharmacologic treatments in treatment-resistant patients or in patients with comorbidities that would have excluded them from conventional randomized controlled trials. Additional research is needed to determine whether N-of-1 trials are suitable for improving the selection of depression treatments in clinical practice.

Keywords

depression; N-of-1; personalized medicine

Corresponding Author: Ian M. Kronish, MD, MPH, Columbia University Medical Center, 622 West 168th Street, PH9-311, New York, NY 10032 Phone: 212-342-1335; Fax 212-342-3431; ik2293@columbia.edu.

Conflicts of Interest

The authors have no conflicts of interest to report.

Introduction

Major depression is one of the most common psychiatric conditions, affecting approximately one in six individuals over the course of their lifetime.¹ Depression is also a leading cause of disability worldwide.² Depression increases risk for suicide, impairs quality of life, and decreases work productivity.^{3–5} Unfortunately, approximately 40% of patients do not respond to first-line depression treatments such as antidepressants or psychotherapy.^{6,7} While specific antidepressants have similar efficacy on average, there are large inter-individual differences in how patients respond to specific depression treatments, both with respect to benefits and harms.⁸ Many patients drop out of treatment after an initial treatment failure.⁹ Yet, studies show that up to 50% of patients who fail to respond to a first depression treatment can obtain benefit after switching to a second treatment.¹⁰ This suggests that improving the selection of the initial depression treatment can greatly benefit patients suffering from depression.

Personalized medicine seeks to optimize the selection of treatments based on a patient's personal characteristics. In the field of depression, despite some promising data, there are few examples of personal characteristics or biomarkers that can reliably predict treatment responsiveness at the individual level. For example, although cytochrome P450 system 2D6 polymorphisms have been associated with poor metabolism and side effects from antidepressants among some patients, evidence for routinely screening for these polymorphisms remains lacking.^{11,12} More recently, brain positive emission tomography (PET) neuroimaging findings have predicted whether patients benefit from antidepressant versus cognitive behavioral therapy.¹³ Yet, this finding remains to be replicated or applied to clinical settings.¹⁴ A recent systematic review seeking to identify characteristics that could be used to predict treatment response concluded that there were few robust predictors available.¹⁵

Another potential approach to personalizing depression treatments is the use of personalized trials, commonly referred to as N-of-1 trials in the research literature. N-of-1 trials lie within the family of single case design studies.¹⁶ The defining characteristic of an N-of-1 trial is the prospective crossover design (e.g., A-B-A-B) within an individual participant in which one period (A) is the treatment being studied, and another period (B) is the treatment being compared.¹⁷ The use of *multiple* crossovers increases confidence in the reliability of results, and is required feature of N-of-1 trials according to some expert classifications.¹⁷ Yet, even simple A-B crossover designs are considered N-of-1 trials by other experts so long as the primary goal is to inform treatment selection among individual patients.¹⁸ Another key characteristic of N-of-1 trials is the systematic collection of data on treatment effects including patient-important outcomes. While the primary focus of N-of-1 trials is on the individual patient, data from a series of N-of-1 trials can be pooled to generate an understanding of population-level treatment effects.^{19–21}

N-of-1 trials are ideally suited to clinical problems for which there is uncertainty or clinical equipoise about the best treatment for an individual patient and for which there are reliable measures for assessing treatment effects. Further, they are well suited to comparing treatments such as antidepressants that are expected to have heterogeneous effects both in

terms of benefits and side-effects among different patients such that there is value in learning which treatment is best for the individual. Finally, N-of-1 trials are suitable for chronic conditions that have readily measurable symptoms (See Table 1 for a summary of conditions under which N-of-1 trials are suitable).²²

N-of-1 trials differ from the usual “trial-of-therapy” approach that is the mainstay of clinical practice. In the “trial of therapy” approach, patients are started on a treatment, and the patient’s status is compared from before to after starting the treatment in an informal manner. This approach is susceptible to a biased understanding of treatment effects that are due to recall bias, expectancy effects, and time effects. N-of-1 trials improve upon the “trial of therapy” approach in that patients compare two or more treatments at the outset (or a treatment with a placebo), data on treatment effects are systematically collected, and data are rigorously analyzed to empirically inform the selection of treatments (Figure 1).²³ Rigorous N-of-1 trials can also involve randomization of the treatment sequence and blinding of treatments. These design features can be used to minimize expectancy effects that arise when treatments are provided open label. Balanced treatment sequences (e.g., A-B-B-A) can be used in place of randomization to reduce the potential for time-effects to lead to a biased understanding of treatment effects. Based on these strengths, N-of-1 trials are considered to be among the strongest designs for making individual decisions about a patient’s treatment selection, and are considered level 1 evidence by the Oxford Center for Evidence Based Medicine.²⁴ The prototypical N-of-1 trial is also distinct from conventional research designs in that the treatment options and monitoring parameters can be customized according to patient preferences, and results can be shared with patients at the end of the trial to maximize shared decision-making pertaining to treatment selection.

Depression may be a suitable candidate for personalized trials, as there is substantial heterogeneity of treatment effects for many established depression treatments and uncertainty about which treatment is best for each patient; there are also emerging and off-label depression treatments that have potential benefit for individual patients but lack large randomized clinical trials across many patient populations; depressive symptoms can be assessed regularly with valid tools; and depression is frequently chronic and slowly progressing. On the other hand, the suitability of N-of-1 trials for depression may be limited as some depression treatments have a relatively long onset-of-action (e.g., serotonin-specific receptor inhibitors can take weeks to achieve maximal onset of action) and other others are intended to have irreversible or long-lasting effects (e.g., insight-oriented psychotherapy).

To determine if the personalized, N-of-1 trial approach would be a useful alternative to the current predominant trial-of-therapy approach to selecting depression treatments, we conducted a systematic review of N-of-1 trials for depression. A prior systematic review reviewed the characteristics and treatment implications of N-of-1 trials published in the medical literature.¹⁸ This review, however, was conducted in 2010, did not include search terms targeted to identifying N-of-1 trials for depression, and did not assess study design characteristics relevant to assessing the feasibility of N-of-1 trials for depression. Further, the search strategy limited itself to N-of-1 trials with randomized treatment assignments and may have excluded N-of-1 trials with non-randomized, balanced sequence designs. We aimed to determine the number, types, and quality of N-of-1 trials relevant to depression.

Review of these outcomes was intended to provide clinicians and researchers with the information needed to determine the feasibility and suitability of pursuing N-of-1 trials to personalize the selection of depression treatments for individual patients.

Materials and Methods

The protocol for this systematic review was registered in PROSPERO, a publicly available international database of prospectively registered systematic reviews, prior to conducting the review.²⁵ The reporting of this review conforms to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines.²⁶ This study was supported by funding from the National Institutes of Health and the Patient-Centered Outcomes Research Institute.

Study selection

Studies were eligible for inclusion if they met the following criteria: (1) Population: humans with elevated depressive symptoms; (2) Interventions: all pharmacological and non-pharmacological treatments for depression (i.e., no restrictions on interventions); (3) Types of Studies: single case research designs that involved at least one cross-over between Treatment A and Treatment B; (4) Comparator/control: placebo or active treatment control; and (5) Outcome: depressive symptoms had to be included. There was no requirement for interventions to be administered in a blinded or randomized fashion, nor to include multiple crossovers. Studies were excluded if they did not contain sufficient design detail to determine eligibility (i.e., they consisted primarily of methods and review without presentation of any data or results from an N-of-1 trial) and/or were not available in English.

Data sources

Potentially relevant articles were identified by searching the following biomedical electronic databases from database inception to October 11, 2016: Ovid MEDLINE, EMBASE, the Cochrane Library (all databases), CINAHL, and PsycINFO. All relevant subject headings and free-text terms were used to represent N-of-1 randomized trials and depression. Terms for MEDLINE included: n-of-1.tw OR ((individual or single) adj (patient\$ or participant\$ or subject\$ or case\$)).tw. OR ipd.tw AND exp depressive disorders/ OR Depression/ OR anhedonia/ OR (depress\$ or anhedoni\$).tw. Additional terms were applied to identify clinical trials. These terms were adapted for the other databases. Ongoing studies were also sought through [Clinicaltrials.gov](https://www.clinicaltrials.gov) and the WHO International Clinical Trials Registry Platform. Additional records were identified by scanning the reference lists of relevant studies and reviews, by employing the Similar Articles feature in PubMed, and by using the Cited Reference Search in Scopus.

Study selection

Two reviewers (MH, LF) independently screened titles and abstracts of all the retrieved bibliographic records. Full texts of potentially eligible records passing the title and abstract screening level were retrieved and examined independently by the two reviewers according to the above-mentioned eligibility criteria. A third reviewer (IK) adjudicated disagreements at both screening levels (title/abstract and full text.)

Data extraction

Two investigators (MH, BK) extracted all data, with disagreements again resolved through consensus with a third investigator (IK) present. The following information was extracted, if present: publication year; country where N-of-1 trial was conducted; trial funding; IRB approval; trial setting; rationale for using an N-of-1 approach; participant eligibility criteria and characteristics; trial design including treatments compared, treatment sequence, duration, randomization, allocation, blinding, and use of washout periods; depressive symptom measures and frequency of measurement; method of analysis including sample size determination, responder definition; N-of-1 trial results (number of participants enrolled; number and sequence of periods completed; losses or exclusion of participations after treatment assignment; number of periods analyzed; number of trials for which data were synthesized), and impact of N-of-1 trial on subsequent treatment. Finally, study quality was rated as low or high using criteria established by the CONSORT Extension for N-of-1 Trials (CENT).²⁷

Analysis

The findings of the various studies are summarized in Table 2. As the studies did not report on patient or clinician satisfaction with the N-of-1 trial design, and as treatment comparisons and study populations were not similar across trials, we did not pool results between studies.

Role of the funding source

The funding source had no role in the design; in the collection, analysis, or interpretation of data; nor in the writing of the report or decision to submit for publication.

Results

Of the 2,588 thousand non-duplicate articles identified by our search, 5 met our full inclusion criteria.^{28–32} One study compared the effectiveness of an antidepressant (amitriptyline) with placebo in a patient with multiple somatic symptoms, but this study was excluded as the patient was not being treated for depression and depressive symptoms were not assessed during the trial.³³ Another study compared imipramine with placebo in a patient with intellectual disability with symptoms of distress, but was excluded as depressive symptoms were not directly assessed.³⁴ The study selection process with reasons for exclusions is presented in Figure 2.

Trial setting

Articles were published from 1986 to 2016, and were conducted in the US, Germany, Australia, and the Netherlands. Trial settings included nursing homes, foster care with on-site nurse, inpatient psychiatry unit, and outpatient psychiatry clinics.

Rationale for N-of-1 Trials

In one study, the goal was to determine the compare the effect of d-amphetamine with methylphenidate and to determine the extent of heterogeneity of treatment effect in a series of patients.²⁹ In another study, the goal was to personalize the selection of a treatment

(methylphenidate) in a special population (depressed geriatric nursing home patients) that was excluded from prior randomized trials and in whom there were expectations of substantial inter-individual differences in harms and benefits.²⁸ In one study, the goal was to compare the effectiveness of an off-label use of an antipsychotic (sulpiride) with placebo for treatment of chronic depression, and to determine the extent of heterogeneity of treatment effect across a series of patients.³¹ In the next study, the goal was determine the best dose of an off-label treatment with placebo as well as to determine the extent of heterogeneity between patients in a series of patients.³⁰ In the final study, the goal was to compare the effectiveness of a pharmacologic treatment with placebo in a patient with human immunodeficiency virus (HIV)-associated encephalopathy likely to have been excluded from conventional randomized controlled trials (RCTs).³²

Ethics and funding support

IRB approval was explicitly reported as being obtained for two of the N-of-1 studies; informed consent without explicitly mentioning IRB review was mentioned in two others; and one study did not comment on informed consent. Two studies acknowledged funding support, one with pharmaceutical industry support, and one with combined pharmaceutical and public support.

Participants

Studies enrolled between 1 and 18 patients in their N-of-1 protocol. There was significant heterogeneity in patient characteristics. One study enrolled a single cognitively impaired, depressed patient with HIV-encephalopathy.³² Another study enrolled 3 nursing home-residing geriatric patients, aged 78 to 81 years old.²⁸ Another study enrolled 18 depressed patients, aged 22 to 45 years, admitted to an inpatient psychiatry unit.²⁹ Two studies enrolled treatment-resistant depressed patients with insufficient therapeutic response to at least one trial of an antidepressant during the current depressive episode (n = 15 and n = 10).^{30,31} Only one of the studies explicitly reported how potential participants were recruited, and none reported how many declined to participate after learning about the study protocol.³⁰

Trial Design Features

Active treatments assessed in the studies included methylphenidate, d-amphetamine, ketamine, and sulpiride. Three studies include an inactive placebo.^{28,31,32} One study included a sham treatment (midazolam for comparison with ketamine);³⁰ this study compared different doses of the same treatment to one another and to the sham. All studies were double-blinded. Treatment periods ranged from one day to 6 weeks. Washout periods ranged from less than 1 day to 1 week. The total duration of the N-of-1 trial protocols ranged from 2 days to 28 weeks. Two studies had adaptive protocols that allowed for clinician-directed changes in dosing as part of the protocol;^{31,32} dosages could be increased or decreased at the discretion of the treating physician who was blinded to treatment assignment. In another study, the N-of-1 trial protocol could be stopped early if participants met remission criteria after a prior treatment dose.³⁰

All studies included randomization, although in one case, this was only with respect to the timing of the insertion of a sham treatment, but not the sequence of increasing doses.³⁰ In

the other studies, randomization was at the level of treatment blocks with balanced sequences. Only one study provided details pertaining to randomization technique and treatment allocation.³⁰ Three studies involved at least one treatment repetition. In one study, a medication was repeated, but at different escalating doses.³⁰

Depressive Symptom Measurement

In two studies, the measure for assessing depressive symptoms was clinician-administered (Montgomery Asberg Depression Rating Scale);^{28,30} in two studies, the depression measure was patient-reported (Hopkins-Symptoms Checklist-depression subscale;^{29,31} mood visual analog scale (VAS) created by the investigators); and in one study, a combination of patient and clinician-administered scales was used (patient- and psychiatrist-administered global drug effect VAS scales and psychiatrist administered Hamilton Depression Ratings Scale).³¹ All studies assessed treatment side effects, although the method for doing so was unclear in one study.

Analytic approach

Quantitative approaches were used to assess for significant inter-individual differences in treatment response in two studies. One used a one-sided paired Student's t-test with an alpha of 0.1 to denote significant difference and the other used the autoregressive integrated moving average (ARIMA) method that accounts for autocorrelations in time series data to compare differences across treatment periods.^{28,31} The remaining studies used cutpoints denoting treatment response or remission to determine individual treatment response but did not use statistical tests of significance to compare individual-level treatment effects.

Drop Outs and Compliance

There were no drop outs in 3 studies;^{28,29,32} however, these studies were all conducted in monitored settings (nursing home, foster home, and inpatient). In one study of self-administered oral medication, 8 of 10 patients completed the study protocol, with 1 patient dropping out early due to perception of clear benefit on active treatment, and 1 patient withdrawing after 13 weeks.³¹ In the remaining study, 9 of 15 patients completed the study protocol, with 6 dropping out prematurely, all of whom had not obtained treatment benefit on the initial dose.³⁰ The three studies that involved self-administered medications did not include a medication adherence check.

N-of-1 Trial Findings

In all 4 studies that enrolled more than one patient, the investigators confirmed that there were individual-level differences in treatment responses that supported the use of the N-of-1 trial approach. (Figure 3) In the ketamine dosing study, the investigators learned that participants responded to treatment at a range of doses from 0.1 mg/kg to 0.4 mg/kg. In the study of depressed geriatric patients, 2 of the 3 participants responded to methylphenidate.³⁰ In the study of depressed inpatients, the investigators learned that 5 participants responded to methylphenidate and d-amphetamine, 1 responded to neither, 7 responded to d-amphetamine, and 5 responded to methylphenidate.²⁹ In the study of treatment-resistant depressed outpatients, 7 of 8 participants responded to sulpiride versus placebo in at least 1

of 2 treatment periods, but only 3 of 8 participants responded to sulpiride in both treatment periods.³¹ Finally, in the study with a single patient, the investigators learned that the active medication was superior to placebo.³² No trials reported on clinician or patient satisfaction with N-of-1 trial design nor included a comparison of the N-of-1 trial approach versus usual care.

Quality rating

None of the studies was rated as high quality according to CONSORT-CENT extension. The study that came closest to being a high quality N-of-1 trial was the one by Maier which included careful consideration of the requisite time blocks needed (i.e., sample size) in terms of individual level measurements, sequence repetition, and used a rigorous statistical approach to compare treatment effects that accounted for autocorrelation in time series data.³¹ Two studies did not include any repetitions of treatment, and hence, did not meet the minimal requirements for an N-of-1 trial according to some definitions.¹⁶

Discussion

Through this systematic review, we aimed to understand the conditions under which N-of-1 trials could be feasible for personalizing depression treatment selection in clinical settings. Accordingly, we identified the range of treatments compared in published N-of-1 trials, the types of populations studied, the methodologies used to gather data, patient acceptability and compliance, and perceived helpfulness to patients and their clinicians. This systematic review demonstrated that it is possible to use the N-of-1 trial approach to personalize the selection of depression treatments in diverse contexts. N-of-1 trials were feasible both in monitored settings where treatments and outcome assessments were administered by clinicians and in outpatient settings where medications and outcome assessments were self-administered. Two of the published studies included patients who would have been likely to be excluded from conventional RCTs. This suggests that N-of-1 trials may be particularly well suited to evaluating treatments in patients with multiple comorbid health conditions who lack evidence on the effects of treatment in patients like them. While the included trials demonstrated that N-of-1 trials are a potentially useful methodology, none of them provided details on implementation challenges such as cost, time, willingness of patients to participate, and perceived usefulness of the methodology to patients and clinicians. These challenges have limited the uptake of N-of-1 trials in clinical practice.^{35,36}

Although most studies only evaluated medications with short onset and washout periods, namely amphetamines and ketamine, one study demonstrated that N-of-1 trials were also feasible for a medication (sulpiride) with duration of onset and washout more similar to second-generation antidepressant medications.³¹ This suggests that there is potential for expanding N-of-1 trials to compare more commonly used first-line antidepressant medications with similar pharmacodynamics properties.

All of the trials sought to double-blind the treatments. In the case of the trial of ketamine, the investigators additionally sought to mask the control by using a sham treatment (midazolam). Of note, there is debate in the field of N-of-1 trials as to the necessity of blinding treatments. Some argue that if the goal is to learn about the best treatment for a

single patient, then blinding may be counterproductive as expectancy effects are an important part of the total treatment effect for an individual patient.³⁷ Further, in a recent study that surveyed patients with chronic diseases about attitudes toward N-of-1 trials, patients disliked the notion of blinding treatments.³⁸ The special packaging and compounding required of blinding medications can also be costly, limiting the feasibility of broadly incorporating N-of-1 trials into clinical practice. In contrast, if the goal of the N-of-1 trial is to isolate the biological effect of an active treatment to a placebo, then blinding will likely be necessary. Ultimately, the decision to blind treatments tested in future N-of-1 trials for depression will depend on the goal of the N-of-1 trialist. While behavioral and psychological treatments were not tested in any of the studies in this review, it is worth highlighting that it is not possible to blind patients receiving such treatments.

The published N-of-1 trials primarily tested off-label or second-line medication treatments that lacked robust data from conventional RCTs. No trials tested complementary and alternative medicine (CAM) treatments (e.g., bright light therapy). Given the rationale of using N-of-1 trials to test treatments that lack conventional RCT evidence, CAM may be a useful class of depression treatments to consider for testing in future N-of-1 trials.³⁹

None of the published N-of-1 trials tested psychological or other behavioral approaches to treating depression (e.g., exercise). In the case of psychological therapies, this was likely due to the presumed lack of reversibility of psychotherapy. Nevertheless, the lack of prior published studies does not negate the potential for psychological therapies to be tested through N-of-1 trials. Although many psychological therapies are designed to be insight-oriented and irreversible, there is evidence that the effect of psychological therapies wane if not maintained. Thus, one could imagine designing N-of-1 trials that compared brief psychological interventions or different maintenance strategies over time. While we found no examples of such behavioral therapies tested in N-of-1 trials for depression, such approaches have been tested in N-of-1 trials for other conditions.^{40,41}

Only one of the included trials compared treatment effects using robust time series analyses that account for autocorrelations between treatment effects across time.³¹ There are now robust statistical techniques that future N-of-1 trialists can apply to their studies.¹⁹ To make such analyses convenient for use in clinical practice, then algorithms may be needed to automate these analyses. Future N-of-1 trials will have to think carefully about how to convey the understanding of these analyses to patients so that they can meaningfully impact on shared-decision making.

There were some important limitations of this review. It is possible that N-of-1 trials were being conducted in clinical settings without publication of data or IRB review. Only one study was conducted in the era of smartphones and mobile health devices, and these devices were not incorporated into the study design.³⁰ In the current era, there may be new opportunities for more robust N-of-1 trials of antidepressant treatments with increased data collection enabled by the use of these technological innovations.³⁸ Finally, the overall conduct and reporting of these published N-of-1 trials did not meet the high quality standards outlined by the CONSORT-CENT extension guidelines. However, these studies were all conducted prior to the publication of these N-of-1 reporting guidelines.

The current evidence suggests that there is potential for N-of-1 trials to be suitable for personalizing depression treatment. Yet, there remains a need for additional studies of N-of-1 trials relevant to depression, including studies that compare the N-of-1 approach with usual care, incorporate measures of patient and clinician satisfaction, and consider issues such as cost-effectiveness and time burden. Developing and sharing methods for conducting N-of-1 trial protocols using mobile health tools may engender progress in this field. Additional questions to be addressed by the field is the necessity for blinding and randomization of treatments, as well as the necessity of multiple crossovers to reduce the potential for a biased understanding of treatment effects. Nevertheless, with advances in mobile health technology and growing interest in personalized treatment selection, these prior successful albeit flawed N-of-1 trials suggest that there may now be new opportunities to design and test N-of-1 trials for depression.

Acknowledgments

Source of Funding

This publication was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1TR001873. This project was also funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services or NIH nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

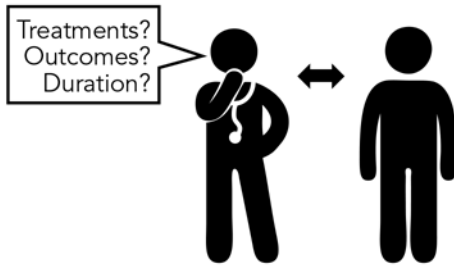
1. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003; 289:3095–3105. [PubMed: 12813115]
2. Unutzer J, Patrick DL, Diehr P, et al. Quality adjusted life years in older adults with depressive symptoms and chronic medical disorders. *Int Psychogeriatr*. 2000; 12:15–33. [PubMed: 10798451]
3. Donohue JM, Pincus HA. Reducing the societal burden of depression: a review of economic costs, quality of care and effects of treatment. *Pharmacoeconomics*. 2007; 25:7–24. [PubMed: 17192115]
4. Hawton K, van Heeringen K. Suicide. *Lancet*. 2009; 373:1372–1381. [PubMed: 19376453]
5. Spitzer RL, Kroenke K, Linzer M, et al. Health-related quality of life in primary care patients with mental disorders. Results from the PRIME-MD 1000 Study. *JAMA*. 1995; 274:1511–1517. [PubMed: 7474219]
6. Practice guideline for the treatment of patients with major depressive disorder (revision). American Psychiatric Association. *Am J Psychiatry*. 2000; 157:1–45.
7. Gartlehner G, Hansen RA, Morgan LC, et al. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. *Ann Intern Med*. 2011; 155:772–785. [PubMed: 22147715]
8. Simon G. Choosing a first-line antidepressant: equal on average does not mean equal for everyone. *JAMA*. 2001; 286:3003–3004. [PubMed: 11743843]
9. Simon GE, Von Korff M, Rutter CM, et al. Treatment process and outcomes for managed care patients receiving new antidepressant prescriptions from psychiatrists and primary care physicians. *Arch Gen Psychiatry*. 2001; 58:395–401. [PubMed: 11296101]
10. Rush AJ, Trivedi MH, Wisniewski SR, et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. *N Engl J Med*. 2006; 354:1231–1242. [PubMed: 16554525]
11. Ozomaro U, Wahlestedt C, Nemeroff CB. Personalized medicine in psychiatry: problems and promises. *BMC Med*. 2013; 11:132. [PubMed: 23680237]

12. Thakur M, Grossman I, McCrory DC, et al. Review of evidence for genetic testing for CYP450 polymorphisms in management of patients with nonpsychotic depression with selective serotonin reuptake inhibitors. *Genet Med*. 2007; 9:826–835. [PubMed: 18091432]
13. McGrath CL, Kelley ME, Holtzheimer PE, et al. Toward a neuroimaging treatment selection biomarker for major depressive disorder. *JAMA Psychiatry*. 2013; 70:821–829. [PubMed: 23760393]
14. Ball TM, Stein MB, Paulus MP. Toward the application of functional neuroimaging to individualized treatment for anxiety and depression. *Depress Anxiety*. 2014; 31:920–933. [PubMed: 25407582]
15. Cuijpers P, Reynolds CF 3rd, Donker T, et al. Personalized treatment of adult depression: medication, psychotherapy, or both? A systematic review. *Depress Anxiety*. 2012; 29:855–864. [PubMed: 22815247]
16. Tate RL, Perdices M, Rosenkoetter U, et al. The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *J Clin Epidemiol*. 2016; 73:142–152. [PubMed: 27101888]
17. Shamseer L, Sampson M, Bukutu C, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015: explanation and elaboration. *J Clin Epidemiol*. 2016; 76:18–46. [PubMed: 26272791]
18. Gabler NB, Duan N, Vohra S, et al. N-of-1 trials in the medical literature: a systematic review. *Med Care*. 2011; 49:761–768. [PubMed: 21478771]
19. Rochon J. A statistical model for the “N-of-1” study. *J Clin Epidemiol*. 1990; 43:499–508. [PubMed: 2139111]
20. Zucker DR, Ruthazer R, Schmid CH. Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: methodologic considerations. *J Clin Epidemiol*. 2010; 63:1312–1323. [PubMed: 20863658]
21. Zucker DR, Schmid CH, McIntosh MW, et al. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *J Clin Epidemiol*. 1997; 50:401–410. [PubMed: 9179098]
22. Kravitz, RL, Duan, N., the DEcIDE Methods Center N-of-1 Guidance Panel (Duan N EI, Gabler NB, Kaplan HC, Kravitz RL, Larson EB, Pace WD, Schmid CH, Sim I, Vohra S)., editors. Design and Implementation of N-of-1 Trials: A User’s Guide. Rockville, MD: Agency for Healthcare Research and Quality; 2014. AHRQ Publication No. 13(14)-EHC122-EF
23. Mirza RD, Punja S, Vohra S, et al. The history and development of N-of-1 trials. *J R Soc Med*. 2017; 110:330–340. [PubMed: 28776473]
24. OCEBM Levels of Evidence Working Group. The Oxford 2011 Levels of Evidence. Oxford: Centre for Evidence-Based Medicine; <http://www.cebm.net/index.aspx?o=5653>. Date accessed: December 21, 2017
25. Kronish, I., Davidson, K., Falzon, L., Hampsey, M. N-of-1 trials for depressive symptoms: a systematic review. PROSPERO. Available at: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42015027574. Date accessed: December 21, 2017
26. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015; 4:1. [PubMed: 25554246]
27. Vohra S, Shamseer L, Sampson M, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ*. 2015; 350:h1738. [PubMed: 25976398]
28. Jansen IH, Olde Rikkert MG, Hulsbos HA, et al. Toward individualized evidence-based medicine: five “N of 1” trials of methylphenidate in geriatric patients. *J Am Geriatr Soc*. 2001; 49:474–476.
29. Little KY. d-Amphetamine versus methylphenidate effects in depressed inpatients. *J Clin Psychiatry*. 1993; 54:349–355. [PubMed: 8407854]
30. Loo CK, Galvez V, O’Keefe E, et al. Placebo-controlled pilot trial testing dose titration and intravenous, intramuscular and subcutaneous routes for ketamine in depression. *Acta Psychiatr Scand*. 2016; 134:48–56. [PubMed: 27028832]
31. Maier W, Benkert O. Treatment of chronic depression with sulpiride: evidence of efficacy in placebo-controlled single case studies. *Psychopharmacology*. 1994; 115:495–501. [PubMed: 7871094]

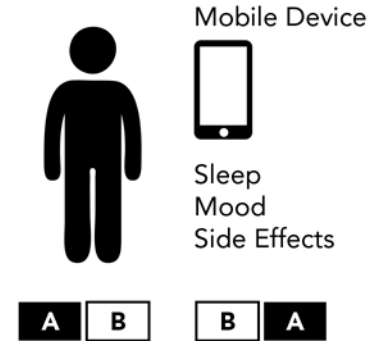
32. White JC, Christensen JF, Singer CM. Methylphenidate as a treatment for depression in acquired immunodeficiency syndrome: an n-of-1 trial. *J Clin Psychiatry*. 1992; 53:153–156. [PubMed: 1592840]
33. Cook DJ, Guyatt GH, Davis C, et al. A diagnostic and therapeutic N-of-1 randomized trial. *Can J Psychiatry*. 1993; 38:251–254. [PubMed: 8518976]
34. Field CJ, Aman MG, White AJ, et al. A single-subject study of imipramine in a mentally retarded woman with depressive symptoms. *J Ment Defic Res*. 1986; 30:191–198. [PubMed: 3735414]
35. Kravitz RL, Duan N, Niedzinski EJ, et al. What ever happened to N-of-1 trials? Insiders' perspectives and a look to the future. *Milbank Q*. 2008; 86:533–555. [PubMed: 19120979]
36. Kravitz RL, Paterniti DA, Hay MC, et al. Marketing therapeutic precision: Potential facilitators and barriers to adoption of n-of-1 trials. *Contemp Clin Trials*. 2009; 30:436–445. [PubMed: 19375521]
37. Schmid CH, Trikalinos TA, Olkin I. Bayesian network meta-analysis for unordered categorical outcomes with incomplete data. *Res Synth Methods*. 2014; 5:162–185. [PubMed: 26052655]
38. Kronish IM, Alcantara C, Duer-Hefele J, et al. Patients and primary care providers identify opportunities for personalized (N-of-1) trials in the mobile health era. *J Clin Epidemiol*. 2017; 89:236–237. [PubMed: 28648833]
39. Johnston BC, Mills E. n-of-1 randomized controlled trials: an opportunity for complementary and alternative medicine evaluation. *J Altern Complement Med*. 2004; 10:979–984. [PubMed: 15673992]
40. Pelham WE Jr, Carlson C, Sams SE, et al. Separate and combined effects of methylphenidate and behavior modification on boys with attention deficit-hyperactivity disorder in the classroom. *J Consult Clin Psychol*. 1993; 61:506–515. [PubMed: 8326053]
41. Sniehotta FF, Presseau J, Hobbs N, et al. Testing self-regulation interventions to increase walking using factorial randomized N-of-1 trials. *Health Psychol*. 2012; 31:733–737. [PubMed: 22353025]

N OF 1 PROTOTYPE

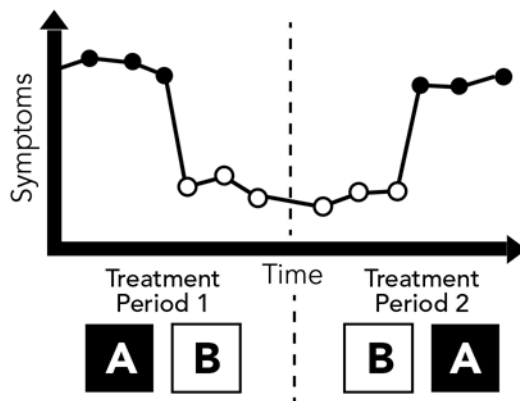
1 N of 1 Design



2 Treatment Tracking



3 Data Visualization & Analysis



4 Shared Decision Making



Figure 1. Prototypical N-of-1 Trial

Step 1 involves customizing the N-of-1 trial according to patient and clinician preferences. Step 2 involves patients tracking treatment effects via diaries or mobile health devices. Step 3 involves statistically analyzing and visualizing results. Step 4 involves treatment selection via shared decision making informed by patient data.

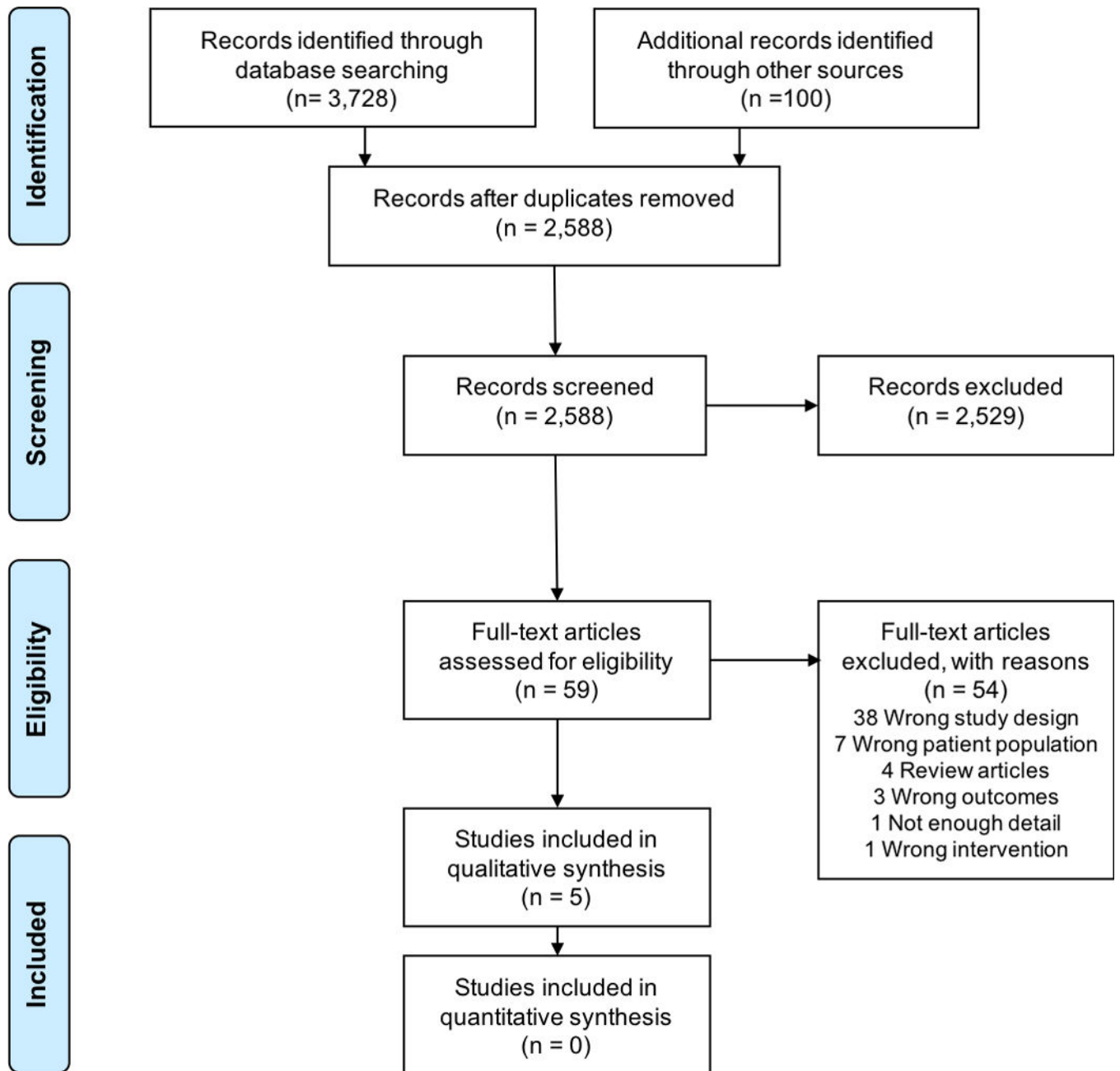


Figure 2.
Flow Diagram of Study Selection











| | | | |
|-----------------------|--|---|--|
| White 1992 (N=1) | Methylphenidate preferred  | No difference | Placebo preferred |
| Little 1993 (N=18) | Methylphenidate preferred  | No difference  | d-Amphetamine preferred  |
| Maier 1994 (N=8) | Sulpiride preferred  | No difference  | Placebo preferred |
| Jansen 2001 (N=3) | Methylphenidate preferred  | No difference | Placebo preferred  |
| Loo 2016 (N=15) | Ketamine preferred  | No difference  | Midazolam preferred |

Figure 3. Outcomes of Published N-of-1 Trials of Depression Treatments

Table 1

Criteria for Determining Whether a Health Condition is Suitable for N-of-1 Trials in Clinical Practice

| N-of-1 trials are suitable if the following conditions are satisfied: | |
|---|---|
| 1. Nature of the Problem | <ul style="list-style-type: none"> • Chronic stable, slowly progressive, or frequently recurring • Symptomatic conditions or asymptomatic conditions with biomarkers that can be tracked over time |
| 2. Nature of the Treatment | <ul style="list-style-type: none"> • Uncertainty about best treatment • Substantial differences in individual responses to treatment • Rapid onset of action • Rapid and safe washout |
| 3. Outcome Assessment | <ul style="list-style-type: none"> • Availability of valid, repeatable measures of treatment effects |
| 4. Stakeholders | <ul style="list-style-type: none"> • Patients, healthcare providers, and health system willing to engage in N-of-1 trials |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Characteristics of Published N-of-1 Trials for Depression*

| First Author, Year Published, Country | Number of Patients | Patient Characteristics | Rationale for N-of-1 Design | Treatments, Blinding | Treatment Sequence, Randomization, Washout, and Duration | Depressive Symptom and Other Outcome Measures | Analytic Approach |
|---------------------------------------|--------------------|---|---|--|--|---|---|
| White ³² , 1992, US | 1 | Foster care residing depressed adult with HIV-associated encephalopathy, resistant to TCA treatment | Determine effectiveness in a patient with a comorbidity that would have made him ineligible for most conventional RCTs | Methylphenidate vs placebo, double-blinded | Counterbalanced, randomized block design: ABA or BAB; number of blocks: 1; treatment period duration: 2 weeks; washout period: <1 day; trial duration: 6 weeks | Patient-reported mood VAS assessed twice per day and physician-reported HAM-D, digit span, symbol digit substitution test, side-effects, nurse assessed sleep and mood as per chart | Positive change in symptoms on outcomes scales, no statistical testing |
| Little ²⁹ , 1993, US | 18 | Depressed inpatients, age 24-45 years, no antidepressant medication use within one week | Determine if there is clinically significant heterogeneity of treatment effects of d-amphetamine vs methylphenidate in depressed patients | d-Amphetamine vs methylphenidate, double-blinded | Counterbalanced, randomized block design: AB or BA; number of blocks: 1; treatment period: 1 day; washout: <1 day; trial duration: 2 days | Patient completed VAS appetite scale, suicidality scale, memory tests, psychiatrist-completed the HAM-D, psychiatrist and patient rated global drug effect VAS scale (worse, no change, better) | Positive change in symptoms on outcomes scales, no statistical testing |
| Maier ³¹ , 1994, Germany | 10 | Chronically depressed community-dwelling adults (>2 years duration), at least one 6-week trial with tricyclic without success within 3 months | Gain preliminary evidence for an untested treatment for chronic, treatment resistant depression, and to determine the extent of inter-individual differences in treatment effects | Sulpiride vs placebo, double-blinded | Counterbalanced, randomized block design (ABAB or BABA); number of blocks: 1; treatment period: 6 weeks; washout: 1 week; trial duration: 28 weeks | Daily evening patient ratings using Hopkins Checklist 90 (3 subscales: depression, anxiety, and somatization as a proxy for drug side-effects) | ARIMA method used to compare sulpiride versus placebo in each pair of treatment periods |
| Jansen ²⁸ , 2001, Holland | 3 | Depressed geriatric patients residing in nursing homes | Guide treatment decisions regarding a controversial medication with uncertain individual-level harms and benefits in depressed geriatric patients | Methylphenidate vs placebo, double-blinded | Counterbalanced, randomized block design (AB or BA); number of blocks: 5; treatment period: 2 days; washout: 1-2 days; trial duration: 5 weeks | Geriatrician completed MADRS and the Barthel Activities of Daily Living scale after each 2 day treatment period; side effects were assessed through method unclear | Paired t-test with alpha 0.1 to denote significance |

| First Author, Year Published, Country | Number of Patients | Patient Characteristics | Rationale for N-of-1 Design | Treatments, Blinding | Treatment Sequence, Randomization, Washout, and Duration | Depressive Symptom and Other Outcome Measures | Analytic Approach |
|---------------------------------------|--------------------|--|---|---|---|--|--|
| Loo ³⁰ , 2016, Australia | 15 | Treatment refractory depressed adults >4 weeks duration, insufficient response to one or more antidepressant medications | To assess the feasibility, safety, and efficacy of a practical individual dose titration approach | Ketamine at 5 increasing doses vs active placebo (midazolam), administered via IV (N=5), intramuscular (N=5), or subcutaneous (N=5) route, subjects evaluated weekly at drug administration, no dose-specific repetitions | Ascending dose design block with random insertion of sham between the first 3 doses (A ₁ -sham-A ₂ A ₃ A ₄ A ₅); number of blocks: 1; treatment period: 1 week; washout period: <1 day; trial duration: 6 weeks | Observer-administered MADRS at 0.4, 2.4, 7.2, 14.4 hours after each treatment period; psychotomimetic, cognitive and physical side effects including HR, BP assessed before and after each treatment | Positive change in symptoms on outcomes scales, no statistical testing |

Abbreviations: TCA, tricyclic antidepressant; RCT, randomized controlled trial; MADRS, Montgomery Asberg Depression Rating Scale; ARIMA, autoregressive integrated moving average; HAM-D, Hamilton Depression Ratings Scale

* *Period*: The time during which a single treatment (A or B) is administered. The order of periods within a pair or treatment block may be randomized. *Block*: A repeated unit of a set number of periods in N-of-1 trials; the sequence of periods may be randomized (for example, three repeating blocks of four periods may look like "AABB BBA, ABAB"). *Sequence*: Multiple blocks comprise an entire sequence. The sequence is the consecutive set of periods, which may or may not indicate size of the repeated unit. *Washout period*: A period in which no intervention is administered. A washout may be administered between different treatment periods or may act as a period in itself, as in a "reversal" design (to "wash out" the effects of a treatment before it is re-administered).