RESEARCH METHODS & REPORTING

CONSORT extension for reporting N-of-1 trials (CENT) 2015: Explanation and elaboration

OPEN ACCESS

N-of-1 trials are a useful tool for clinicians who want to determine the effectiveness of a treatment in a particular individual. The reporting of N-of-1 trials has been variable and incomplete, hindering their usefulness in clinical decision making and by future researchers. This document presents the CONSORT (Consolidated Standards of Reporting Trials) extension for N-of-1 trials (CENT 2015). CENT 2015 extends the CONSORT 2010 guidance to facilitate the preparation and appraisal of reports of an individual N-of-1 trial or a series of prospectively planned, multiple, crossover N-of-1 trials. CENT 2015 elaborates on 14 items of the CONSORT 2010 checklist, totalling 25 checklist items (44 sub-items), and recommends diagrams to help authors document the progress of one participant through a trial or more than one participant through a trial or series of trials, as applicable. Examples of good reporting and evidence based rationale for CENT 2015 checklist items are provided.

Larissa Shamseer *PhD candidate and senior research associate*[1], Margaret Sampson *manager*[2], Cecilia Bukutu *associate director*[3], Christopher H Schmid *professor of biostatistics*[4], Jane Nikles *NHMRC postdoctoral research fellow*[5], Robyn Tate *professorial research fellow*[6], Bradley C Johnston *assistant professor*[7], Deborah Zucker *adjunct assistant professor*[8], William R Shadish *professor*[9], Richard Kravitz *professor and co-vice chair of research*[10], Gordon Guyatt *professor*[11], Douglas G Altman *professor*[12], David Moher *senior scientist*[1], Sunita Vohra *centennial professor*[13], and the CENT group

[1]Clinical Epidemiology Program, Ottawa Hospital Research Institute; University of Ottawa, Canada; [2]Library Services, Children's Hospital of Eastern Ontario, Canada; [3]Child and Youth Data Laboratory, Alberta Centre for Child, Family and Community Research, Canada; [4]Department of Biostatistics and Center for Evidence Based Medicine, Brown University, USA; [5]University of Queensland, Australia; [6]Centre for Rehabilitation Research, Sydney Medical School - Northern, University of Sydney, Australia; [7]Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, University of Toronto, Canada; Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Canada; [8]Tufts University School of Medicine, USA; [9]University of California, Merced, USA; [10]Department of Internal Medicine, University of California, Davis, USA; [11]Department of Clinical Epidemiology & Biostatistics, McMaster University, Canada; [12]Centre for Statistics in Medicine, University of Oxford, UK; [13]Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, Canada

The ultimate stakeholder in medical decision making is the patient. Although parallel group, randomised controlled trials (RCTs) are the gold standard for generating evidence about group treatment efficacy, such evidence is not always available or relevant for individual treatment decisions. Many scenarios exist where this holds true. As examples, unique challenges exist with respect to evaluating treatments in populations with rare diseases; adequate recruitment for group trials is not always feasible, retention may be difficult, and funding may be difficult to obtain.[1 2] Similarly, children, adolescents, and elderly people are typically excluded from or not studied in large scale RCTs.[1 3]

Patients with comorbid conditions or receiving concurrent therapies are also understudied because of stringent inclusion protocols in RCTs. Exclusions of these groups may occur for safety reasons, but more often exclusions according to strict eligibility criteria occur in order to ensure a homogenous sample in the hope of increasing the likelihood of demonstrating a treatment effect.[4 5] Unfortunately, the patient groups described above comprise the majority, rather than a subset, of the clinical population, and information to guide their treatment decisions is sparse.[3 6 7]

Correspondence to: S Vohra svohra@ualberta.ca

Difficulty in measuring and accounting for heterogeneity provides another reason why group trials may not always be the best choice of study design. Summary data from group trials likely contains some level of heterogeneity and may not predict an individual's response to treatment. Even when there is a clear overall benefit at the group level, it may not benefit individual patients equally, or at all (fig 1⇓). N-of-1 trials provide a mechanism to evaluate the effects of treatment for an individual; when trials of individuals are combined using the right statistical techniques, they may be able to approximate effect estimates from group data.[8]

In the context of making decisions about an individual patient's care, N-of-1 trials have been considered to be among the most relevant and rigorous study designs for assessing treatment efficacy; they are listed as "level 1" evidence in the Oxford Centre for Evidence-Based Medicine 2011 levels of evidence.[9 10] As with crossover trials, N-of-1 trials eliminate confounding by covariates since each patient serves as his or her own control. The use of multiple crossovers within well designed N-of-1 trials[11] increases confidence in the reliability of the results.

In addition to their value in evidence based clinical practice, N-of-1 trials also have a valuable role in advancing medical research evidence. For example, researchers might conduct a series of N-of-1 trials to inform overall treatment effect for a group, while simultaneously obtaining relevant treatment information for individual participants. Furthermore, N-of-1 trials might be useful in personalised medicine research to explore subgroups that have differing responses to treatment, complementing the emerging field of pharmacogenomics research. If done and reported well, N-of-1 trials can make a worthwhile contribution to patient centred research, as they can empower patients to participate actively in selecting their treatment options.

Poor reporting, labelling, and indexing of N-of-1 trials has, to date, prevented an accurate estimate of the prevalence of trials in the literature. None the less, N-of-1 trials have been documented evaluating a range of health conditions, including mental and behavioural disorders and diseases of the nervous, respiratory, circulatory, musculoskeletal, and digestive systems.[12] They have also been used to evaluate a variety of interventions, whether pharmacological or non-pharmacological, including complementary or alternative therapies. Most documented trials are being done in Western regions (North America, Europe, and Australia).

Defining "N-of-1" trials

The term "N-of-1 trial" is shared between the fields of medicine and behavioural science, but refers to different concepts within each.

N-of-1 trials are a well established and extensively used experimental design in the behavioural sciences.[13-15] The term is often used to refer to a range of single case experimental designs (fig 2⇓).[16 17]

In medicine, "N-of-1" largely refers to a specific trial design—one using a repeated cycle of treatment challenge and withdrawal (A-B-A-B) in a single participant, where one period ("A") is the treatment being studied and the other period ("B") is a comparison treatment, a control, or no intervention.[18] This design is sometimes described as "ABAB" and may incorporate key elements of RCTs used to reduce bias such as randomisation (of treatment sequence) and blinding (of patient, care provider, outcome assessor, and data analysts). In the remainder of this document, the term "N-of-1 trial" will refer to a prospective, multiple crossover ABAB, single participant trial used in

medicine. Terms often used to describe methodological aspects of N-of-1 trials are provided in box 1.

Evidence of incomplete and inaccurate reporting

The reporting of key elements of N-of-1 trials varies, as characterised in a recent systematic review of N-of-1 trials examining health interventions for medical or clinical conditions, identifying 100 reports for inclusion: 60 series of several N-of-1 trials and 40 individual trials.[12] Although randomisation is not essential in N-of-1 trials, trials that were labelled as randomised (n=71) described the methods of sequence generation only 30% of the time and failed to indicate whether allocation concealment was used 76% of the time. Perhaps more concerning is that a primary outcome was not indicated in 79% of included reports, and 64% of reports did not state whether harms had occurred. Most trials reported the use of statistical analyses (n=75), and, of these, 89% reported summary measures, yet only 49% provided estimates of precision. Reporting of other important trial characteristics that may lead to bias in the interpretation of results, such as whether carryover effect or period effect were assessed, were not reported in 91% and 97% of trials, respectively.

These findings are in line with an earlier systematic review of 108 medical N-of-1 trials, in which less than half (45%) reported enough information to enable meta-analysis; specific missing elements were measures of variance and precision.[19] Furthermore, 47% of trials did not provide any numerical estimate of effect size.

Failure to report key elements of the methods and results for N-of-1 trials impedes readers' assessment of the validity of the research and prevents clinicians and researchers from making optimal use of N-of-1 trial findings in clinical care and future research. The CENT 2015 guidance is in line with the recent international efforts towards better reporting of health research overall,[20] and authors are encouraged to make use of it when preparing their reports of individual or series of N-of-1 trials. Likewise, those charged with reviewing N-of-1 trials for publication (that is, made publicly available in some form) are urged to use CENT to ensure that information is complete before acceptance.

Scope of CENT 2015

The CENT 2015 guidance is aimed at facilitating the reporting of primary N-of-1 trials—individual trials and prospective series of N-of-1 trials. It is not, however, intended to address the reporting of retrospective syntheses (that is, systematic reviews or meta-analyses) of data from separate N-of-1 trial reports. Box 2 clarifies the distinction between primary and secondary studies. CENT is also not intended for other single case experimental design studies used in behavioural medicine. Additional guidance for these types of studies is under way[21] or planned by members of the CENT group.

CONSORT 2010

The CONSORT (Consolidated Standards Of Reporting Trials) Statement was among the first consensus based reporting guidelines to appear in the mid-1990s and has since been updated, most recently in 2010, to remain in line with new evidence and opinion.[22] It is intended to provide authors with a minimum set of items that should be addressed in reports of two arm, parallel group, clinical trials. The latest iteration,

Box 1: Methodological terminology typical in N-of-1 trial reports

N-of-1 trial—An experimental clinical study design to determine the effect of an intervention in a single study participant. CENT is intended to be used to report repeated challenge-withdrawal (that is, "ABAB") trials, commonly used in medicine, in which multiple crossovers between treatment(s) and control (placebo, standard care, alternate treatment) are continued for a pre-specified amount of time or until treatment effectiveness is determined. More than two treatment alternatives may be compared to each other or control (that is, "ABCABC")

Period—The time during which a single treatment (A or B) is administered. Period length is typically determined a priori and may vary within a trial. The order of periods within a pair or treatment block may be randomised.

Block or pair—A repeated unit of a set number of period in N-of-1 trials is referred to as a block, in which the sequence of periods may or may not be randomised (for example, three repeating blocks of four periods may look like "AABB BBAA ABAB"). By convention, when the repeated unit contains only two periods (for example, three repeating pairs may look like "AB BA BA"), it is conventionally referred to as a pair.

Sequence—Multiple pairs or 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).

Run-in period—A pre-specified duration of time before a trial begins, during which trial treatments may be initiated (for example, to get to a stable therapeutic dose), to determine potential patient compliance with study regimens, or to allow for washout of a medication(s) a patient may have been taking before the trial.

Box 2: Terminology used to describe primary and secondary reports of N-of-1 trials

N-of-1 trial—A prospective, multiple crossover (that is, ABAB) trial in a single participant.

Series of N-of-1 trials—A prospectively planned set of N-of-1 trials designed to evaluate the same clinical question across individuals. A report of a series of N-of-1 trials may include quantitative synthesis such as meta-analysis.

Systematic review of N-of-1 trials—A systematic collection of N-of-1 trials in a single report using explicit a priori methodology including systematic identification, data collection, and analyses processes. Data from individual trials may be synthesised using narrative or, in certain circumstances, meta-analytic methods.

Quantitative synthesis or meta-analysis of N-of-1 trials—The statistical synthesis of data from more than one N-of-1 trial; may be a component of an N-of-1 series, systematic review, or literature review of N-of-1 trials.

CONSORT 2010, consists of a 25 item checklist (37 items including sub-items) and a flow diagram illustrating how to document participants' flow through a trial. It has received widespread support within the biomedical publishing community (endorsed by over 600 journals), and its endorsement is associated with more completely reported trials.[23]

CENT 2015 is an official extension of the CONSORT 2010 Statement and can be found, along with other extensions, on the CONSORT website (www.consort-statement.org). For journals wishing to endorse CENT 2015, please see the CENT 2015 Statement for recommended text to include in journal "instructions to authors" text.[24]

Overview of checklist development

Selection of candidate checklist items was informed by the aforementioned systematic review[12] and by the CONSORT 2010 checklist. Checklist development followed the general process recommended by the Enhanced Quality and Transparency of Reporting (EQUATOR) Network for developing a reporting guideline,[25] in which consensus is a fundamental component. A two-round Delphi survey of 56 experts—including N-of-1 trialists, epidemiologists, reporting guideline developers, biomedical journal editors, and funders—preceded an in-person consensus meeting of 23 participants held in May 2009. A detailed description of the CENT development process can be found in the CENT Statement.[24] All CENT related guidance documents have undergone an iterative refinement process within the CENT steering committee (DGA, Nick Barrowman, CB, DM, JN, MS, LS, RT, SV) and larger CENT group listed at the end of this document. LS and SV led the writing of this document and members of the CENT group contributed to the writing and identification of relevant examples contained within this document. A subcommittee was also convened to develop the CENT diagrams (NBG, JN, DZ).

CENT 2015 checklist

The CENT 2015 checklist is an extension of 14 items of the 25 CONSORT 2010 items (table 1⇓). Of the 25 items of the CONSORT 2010 checklist, 12 are further divided into sub-items, creating a total 37 sub-items. With the modifications and additions to CONSORT 2010 items, there are 44 sub-items in the CENT checklist, some of which only pertain to series of trials (as indicated). For item 1b, pertaining to the reporting of abstracts, specific recommendations for N-of-1 trials are proposed in table 2⇓.

The recommendations within CENT may require more words and space than N-of-1 trialists are accustomed to. Providing detailed descriptions for some trials will facilitate transparency and future reproducibility, in line with emerging journal policies aimed at facilitating reproducibility.[26]

We recognise that improved reporting must be balanced against patient confidentiality in situations when the condition is rare. Authors must be mindful of this, and if there is any doubt as to whether complete reporting could be potentially identifying, they should seek consultation with their institutional ethics board. This issue is of heightened importance in N-of-1 trials of rare conditions or when the potential societal stigma is high. Caution should be taken when reporting a combination of identifying information pertaining to CENT items 4a, 4b, 14a, and 15.

CENT diagrams

In the spirit of the CONSORT 2010 flow diagram, two diagrams specifically for CENT have been developed to help authors visually depict participant progress and outcomes through an individual trial (fig 3⇓) and the flow of participants in a series of trials (fig 4⇓). We recommend that authors include these diagrams, as appropriate, in reports of N-of-1 trials; specific guidance on the information to include in each is provided in items 17a.1 and 13a.1, respectively.

CENT explanation and examples

In the remainder of this document we provide explanations of each CENT 2015 checklist item with examples of good reporting. While many CENT 2015 items refer directly to CONSORT 2010 items, examples from N-of-1 trials are still provided to give an example of reporting in the context of N-of-1 trials. Where we felt it was necessary, a rationale is provided for specific nuances associated with reporting N-of-1 trials. We have tried to provide examples of reporting from both series and individual N-of-1 trial reports, where applicable and available. For a comprehensive understanding of reporting of an N-of-1 trial, we strongly recommend that this explanatory document be read together with the CONSORT 2010 Explanation and Elaboration document.[27]

As noted, the authors of many N-of-1 trials fail to report essential information, leaving a small pool of studies from which to draw examples of complete reporting.[12] Rather than constructing hypothetical examples of reporting that do not exist in the literature, we rely more heavily on examples from reports of series of trials. For some items, no example of good reporting could be identified. As the CENT guideline becomes established and has an impact on N-of-1 trial reporting, this document will be updated to include a more comprehensive and relevant set of examples.

Finally, for convenience, we will refer to treatments and patients throughout this document, although we recognise that not all interventions evaluated in N-of-1 trials are technically treatments and trial participants are not always patients. All citations within included examples have been removed for ease of reading.

Title and abstract

Item 1a

Standard CONSORT item: Identification as a randomised trial in the title

CENT extension: Identification as an "N-of-1 trial" in the title

Example: "An N-of-1 randomized controlled trial ('N-of-1 trial') of donepezil in the treatment of non-progressive amnestic syndrome"[28]

For series: Identification as a "series of N-of-1 trials" in the title

Example: "Efficacy of temazepam in frequent users: a series of N-of-1 trials"[29]

Explanation: In order for an N-of-1 trial or a series of trials to be easily identified in an electronic database search, the title, at minimum, should contain prominent, recognisable terminology. For instance, the potential for N-of-1 trials to be quantitatively synthesised will largely depend on whether they can be reliably identified in the literature. Including "N-of-1 trial" or "series of N-of-1 trials" in the title of a report will ensure that they are consistently identified, regardless of the capabilities of the search interface. Since N-of-1 trials are currently described using heterogeneous terminology (such as single case experimental study, single patient trials, etc), the use of "N-of-1" is advised to refer to the specific N-of-1 trial design (that is, ABAB) around which CENT is based (that is, prospective, multiple crossover).

Item 1b

Standard CONSORT item: Structured summary of trial design, methods, results, and conclusions

(for specific guidance see CONSORT for abstracts)

CENT extension: see CENT guidance for abstracts (table 2⇓)

Example:

"*Introduction*: There are several substances available which are used for prophylaxis in patients suffering from migraine. To test the effects of second choice drugs (e.g. in case of side effects of first line substances) reliably, n-of-1 trial on a single patient is one viable option. Feverfew (Tanacetum parthenium L.) is a plant of the family of chrysanthemum. It is known as well under 'wrong chamomile' and can be used for prophylaxis of migraine.

Material and Methods: 100 mg extract of feverfew (verum) and 100 mg lactose (placebo) were manufactured in identical caps for a single female patient and taken for two weeks each. After this time, both substances were taken in reverse order. The patient documented on a daily basis whether she had headache and noted the intensity of the pain. The experimental assembly was planned in a double blind design.

Results: During placebo intake, the patient suffered from 3.0 attacks of migraine on a weekly average, while verum (feverfew) reduced the number of attacks to 1.5 per week. During placebo intensity of migraine attacks averaged 3.0 units on a Likert Scale (0 = no pain, 6 = maximal conceivable pain), during verum the respective value was 1.6. The pain intensity was approximately twice as high while taking placebo compared to verum.

Conclusion: Use of feverfew showed clear efficacy in this patient on the basis of an experimental n-of-1 trial. Reduction of attack rate and pain intensity alike was over 50%."[30]

Explanation: While the suggested abstract structure remains the same as CONSORT 2010 (item 1a), there are some differences in content to be considered in abstracts of N-of-1 trial reports. CENT-specific guidance, adapted from the 2008 CONSORT extension for journal and conference abstracts,[31] is proposed in table 2⇓.

Introduction

Item 2a

Standard CONSORT item: Scientific background and explanation of rationale

CENT extension: Replaced by item 2a.1 and 2a.2

Item 2a.1 CENT extension: No change from CONSORT item 2a

Example:

"Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity and mortality worldwide and results in substantial economic and social burden. Many patients with COPD identify dyspnea on exertion as the key adverse impact of their condition and the strongest determinant of their functional status. Several guidelines have suggested that COPD management should be individualized and based on assessments of the impact of treatment on dyspnea rather than on physiological measures of pulmonary function or arterial blood gases.

A Cochrane review has summarized the well recognized mortality benefits of long term oxygen therapy (LTOT) for individuals with chronic resting hypoxemia. Clinicians sometimes prescribe ambulatory oxygen for patients without resting hypoxemia, who experience hypoxemia only during exercise or activities of daily living. Studies of short term ambulatory oxygen have demonstrated improvements in acute

exercise performance among patients with moderate to severe COPD. However, these laboratory based, acute physiological responses may not reflect how a patient responds symptomatically to the longer term use of ambulatory oxygen in their daily lives.

Justifying the expense and inconvenience of long term ambulatory oxygen for transient exercise hypoxemia requires an understanding of its effect on the patient's experience in the community, during activities of daily living. Four randomised controlled trials (RCTs), evaluating the role of ambulatory oxygen, have reported mixed results..."[32]

Explanation: This item emphasises the importance of providing a rationale for the broader study question, unrelated to N-of-1 design. For a detailed explanation, please refer to item 2a in the CONSORT 2010 explanatory document. For explanation of the decision to use N-of-1 trial design please refer to CENT 2015 item 2a.2

Item 2a.2 CENT extension: Rationale for using N-of-1 approach

Example: "Past trials have compared the INR variability of Coumadin with that of generic Barr-warfarin (Barr Laboratories, Pomona, NY) in the US. These clinical studies concluded that the 2 products were interchangeable, although each study assessed INR variation as averaged within groups rather than within individual patients. As of May 2, 2005, no published studies have satisfied all relevant interchangeability concerns: patients as subjects, comparison of variability within and between individuals, and using INR as the outcome. Finding an appropriate design for such a study has been an issue. However, the N-of-1 randomised, crossover study design is useful in making treatment decisions at the individual and group level with small sample sizes...."[33]

Explanation:

It may not be immediately apparent to readers why an N-of-1 trial, rather than a traditional RCT design, was used to study a particular condition, intervention, outcome, or combination of these. Certain patient populations (such as those with rare diseases, paediatric populations) or sub-populations (such as patients with comorbid conditions or those using concurrent therapies) are often overlooked or excluded from study in conventional RCTs. Indeed, RCTs often attempt to limit heterogeneity of the population under study since evaluation of homogenous populations increases the chances of detecting a signal (treatment effect) among the noise (heterogeneity).[34] Therefore, RCT findings may not always be helpful for guiding the treatment approach for a particular patient.

Reasons why an N-of-1 trial is the most appropriate design for evaluation of treatment in a particular patient or set of patients should be stated in the introduction of the N-of-1 trial report. For instance, some conditions better lend themselves to evaluation through N-of-1 design than others. Guyatt et al proposed guidelines for determining when an N-of-1 trial may be appropriate (box 3).[18] It is helpful to report whether these criteria were considered in the decision to carry out the N-of-1 trial(s).

Item 2b

Standard CONSORT item: Specific objectives or hypotheses

CENT extension: No change from CONSORT item 2b

Example: "To describe the use of an N-of-1 randomised clinical trial (N-of-1 RCT) in general practice as illustrated by the case of a 16 year old boy with a learning and attention problem whose parents were convinced that amphetamines were necessary."[35]

Example: "Our objectives were (1) to determine whether in children undergoing doxorubicin-containing chemotherapy, topical vitamin E decreases an objective measurement of oral mucositis compared to placebo and (2) to assess the feasibility of an innovative trial design of combining N-of-1 trials using Bayesian meta-analysis."[36]

Explanation: Same as CONSORT 2010 item 2b.

Methods

Trial design

Item 3a

Standard CONSORT item: Description of trial design (such as parallel, factorial) including allocation ratio

CENT extension: Describe trial design, planned number of periods, and duration of each period (including run-in and washout periods, if applicable)

Example: "Treatment was administered in 3 pairs, each consisting of 2 periods in which either tramadol 50 mg BID or placebo was administered for 6 days, followed by a 2-day washout period, and then the administration of the alternate for 6 days. A 2-day washout period was also carried out after pairs 1 and 2."[37]

In addition for series: Whether and how the design was individualised to each participant, and explanation of the series design

Example: "We considered a wash out period of one week sufficient, given the short half life of the NSAIDs used in our study. Previous n of 1 series and RCTs comparing NSAIDs with paracetamol have used a similar wash out period."[38]

Example: "Each subject underwent an n-of-1 trial. Visit 1 was followed by a two-week run in phase to familiarize participants with the protocol."[39]

Note: No example of good reporting demonstrating individualisation could be identified.

Explanation:

A succinct description of the intended trial design, including number of periods and whether run-in or washout periods were planned, will contribute to the readers' interpretation of the trial. Reporting of specific aspects of trial design is addressed by items 3b through to 13 of the CENT checklist.

A run-in period occurs before a trial begins and is typically used to initiate trial medications (for example, to get to a stable therapeutic dose), determine tolerability, assess potential compliance with study regimens, or to allow for wash-out of medication effects a participant was taking before enrolment in the trial.[40]

A washout period may occur between treatments to allow the effects of one treatment to wear off before proceeding with the

> **Box 3: Guidelines for choosing an N-of-1 trial (reproduced from** *The Users' Guides to the Medical Literature*[9]**)**
>
> Is an N-of-1 trial indicated for this patient?
>
> Is the safety or effectiveness of treatment in doubt?
>
> If effective, will the treatment be continued long term?
>
> Is an N-of-1 trial feasible in this patient?
>
> Is the patient eager to collaborate in the designing and carrying out of an N-of-1 trial?
>
> Does the treatment have rapid onset and termination of action?
>
> Is optimal duration of treatment feasible?
>
> Are targets of treatment that are important to the patient also amenable to measurement?
>
> Can you identify the criteria to end the N-of-1 trial?
>
> Is there a pharmacist who can help?
>
> Are strategies in place for the interpretation of the data?

next (that is, to reduce carryover effect; see CENT item 12c) or it may be used as a trial period in itself in order to allow the effects of the preceding treatment to fade or dissipate (that is, reversal design). A washout may also be incorporated as a part of a treatment. Authors should say which of these roles washout periods are intended to play in the trial design.

Whether run-in or washout periods are employed, it is helpful to give the rationale for their use, and their length should be stated (for example, it takes 'X' days after stopping blood thinning medication for patients' blood coagulation to return to its previous state).

For a series of N-of-1 trials, authors should report any details of trial design (such as period length) that were tailored around a particular participant, describing the individualisations made.

Item 3b

Standard CONSORT item: Important changes to methods after trial commencement, with reasons
CENT extension: No change

Example: "Since we were in the process of model development, the design was modified three times during the study. From pilot II and throughout the study the cimetidine dose was increased to 800 mg, and the patients were asked to register the total duration of their trial. During pilot II the patients were requested to record separate measures for pain, heartburn/acid regurgitations, and global symptoms. The maximum number of doses per day was reduced to two."[41]

Explanation: This item is separate from the notion of intentional changes due to individualisation of trial design (item 3a). It is not possible to predict all possible circumstances in which changes to trial design/methods may be made and whether such changes are warranted. Documentation of any changes that occurred over the course of a trial is encouraged (item 3b). Since N-of-1 trial protocols have previously not been widely available, documenting changes to the trial design is especially important.

Changes to the sequence of periods in a trial occur for different reasons, such as tolerability of treatment, dose interruptions, or even purposeful modification of the order of treatments by participants or physicians. Authors should document any changes that were made, with reasons; this will enable readers' assessment of potential bias in reported findings.

Participant(s)

Item 4a

Standard CONSORT item: Eligibility criteria for participants

CENT extension: Diagnosis/disorder, diagnostic criteria, comorbid conditions, and concurrent therapies

Example: "The participant was a 71-year-old male with the primary diagnosis of dementia of the Alzheimer type. A seizure disorder and a history of congestive heart failure were also present. Medications, excluding study drugs, were Dilantin 300 mg QD, Digoxin 0.25 mg QD, and Milk of Magnesia 30 cc QD on odd days. Haloperidol 1 mg PRN was chosen by the ward physician as back-up for behavioral problems. Haloperidol 1 mg was given on six occasions, all of which were more than 36 hours preceding saliva collection."[42]

For series: same as standard CONSORT item

Example: "There were 43 subjects recruited from respiratory clinics primarily dealing with asthma and COPD in hospital and private practice settings. Inclusion criteria required that subjects were aged between 40 and 80 years, were current smokers or ex-smokers, experienced at least mild shortness of breath on exertion, had a baseline FEV, of < 60% of predicted value and FEV1/forced vital capacity (FVC) ratio <60%, were stable at time of entry into the study (no deteriorations requiring emergency hospital or local medical officer visits or hospital admissions in the previous 28 days), and had poorly reversible airways obstruction defined by the British Thoracic Society (an increase in FEV, of not greater than 15% and 200 mL after salbutamoll'). No attempt was made in our recruitment to distinguish between emphysema and chronic bronchitis. Patients were excluded from entry into the trial if they had a history of asthma by their clinician, unstable airways disease, other respiratory disease, other uncontrolled disease, had changes in their medication in the previous 28 days, or were on beta blocker medication."[39]

Explanation: Since it is rare, in individual trials, for eligibility criteria to be applied, a description of the condition(s) under study should be reported, authors should describe patient characteristics such as the diagnosis, comorbid conditions, and concurrent medications, if relevant. Providing this information will help readers gauge to which populations and subpopulations the findings of trial are applicable. For series of N-of-1 trials, authors should report all of aforementioned details along with specific reasons participants were not eligible for inclusion in the series (that is, exclusion criteria), if applicable.

Item 4b

Standard CONSORT item: Settings and locations where the data were collected

CENT extension: No change

Example: "individuals with insomnia were recruited by three suburban Brisbane general practices, and from the community directly, through regional Queensland newspaper and television media campaigns."[43]

Explanation: See CONSORT 2010 item 4b.

Item 4c

Standard CONSORT item: None (new for CENT)

CENT extension: Whether the trial(s) represents a research study and, if so, whether institutional ethics approval was obtained

Example: "The study was approved by the institutional Research Ethics Board. Each child for whom parental consent was obtained was enrolled in an N-of-1 trial."[44]

Explanation: It has been suggested that an N-of-1 trial undertaken solely to better manage an individual's treatment and that meets clinical ethical standards might be considered a clinical investigation rather than research and so may not require institutional review board oversight.[45 46] The number of N-of-1 trials undertaken for clinical investigation is unknown since accounts of such trials are typically not published.[47] However, investigations around a prospectively designed series of N-of-1 trials intended for comparison or combination in meta-analyses may require research ethical approval, and, if so, authors should indicate this in the study report. If an N-of-1 trial was carried out under the auspices of research[45] authors should clearly state whether a health research ethics board reviewed and approved the research study and whether patient consent was obtained.[48]

Interventions: Item 5

Standard CONSORT item: The interventions for each group with sufficient details to allow replication, including how and when they were actually administered

CENT extension: The interventions for each period with sufficient details to allow replication, including how and when they were actually administered

Example: "A double-blind, randomised, controlled multi-crossover trial consisting of 12 test doses, six with 400 mg cimetidine and six with placebo, was conducted in each patient. To ensure a spread of the cimetidine doses, the test doses were ordered in six pairs, each containing one dose of cimetidine and one dose of placebo. The sequence within each pair was randomised. The patient was instructed to take one test dose when in need of symptomatic relief and to measure its effect within 3-6 h. The patient was also advised to avoid concurrent intake of other alleviating agents or food before the symptomatic effect was measured. A maximum of three doses were allowed per day, with at least 6 h interposed between the doses."[41]

Explanation: A distinguishing feature of N-of-1 trials is that the intervention(s) can generally be tailored to meet a patient's unique profile.[49] Authors should provide the name and content of the intervention(s) as well as the procedures for delivering the treatment(s). We recommend that trial authors consult the TIDieR (Template for Intervention Description and Replication)

checklist for a listing of intervention details that authors should include in their reports.[50] In addition, authors may find the CONSORT extensions for herbal,[51 52] acupuncture,[53] or non-pharmacological[54 55] interventions helpful, if applicable.

Outcomes

Item 6a

Standard CONSORT item: Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed

CENT extension: Replaced by item 6a.1 and 6a.2

Item 6a.1

CENT extension: No change from CONSORT item 6a

Example: "Emetic episodes were recorded by a parent in a study diary. The frequency of emetic episodes was classified by absolute numbers, and categorized into complete response (0 episodes/day), major response (1-2 episodes/day), or failure (>2 episodes/day). The primary outcome was a comparison of the proportion of days in each cycle that patients had a complete response when treated with metopimazine vs. placebo. This outcome was evaluated for a complete cycle of chemotherapy as well as separately for the acute and delayed phases. Secondary outcomes included the absolute number of emetic episodes per day, 'patient distress' as assessed by a parent twice daily (at noon and 8 p.m.) using a 6-face 'happy face' scale ranging from 0 (no distress) to 5 (extreme distress) and the frequency of adverse effects experienced during the treatment and placebo cycles."[44]

Example: "The outcome measures chosen were those which the patient thought were most important, and included well-being; nausea; vomiting; fever; abdominal gas and pain; stool volume, consistency, and odour; and presence of blood in the stool. Throughout the study, the patient kept a diary in which these items were evaluated daily. 10 cm visual analogue scales were used to estimate well-being, nausea, abdominal pain, and gas. Fever, vomiting, and the presence of blood were assessed by means of yes/no response. Stool consistency and odour were measured with a two-point scale (normal/watery for consistency; normal/foul for odour). Stool volume was measured in litres.… Because the patient found stool collection unpleasant, stool volume was recorded only during the second week of each treatment period."[56]

Explanation:

All outcomes measured, whether primary or secondary, should be identified and completely defined. It is well documented that RCTs are selectively reported; for instance, 40-62% of reports of RCTs report a different primary outcome than their protocol.[57] In addition, where outcomes are assessed at more than one time point, the frequency of measurements should be indicated and authors should state if one time point was of primary interest.

Because they are designed around an individual patient, N-of-1 trials allow for collaboration between patients and practitioners. Patients are often involved in this process to help ensure that outcomes important for patients are included.[58-62] In these instances, it is desirable for authors to indicate who selected the outcomes.

Item 6a.2

CENT extension: Description and measurement properties (validity and reliability) of outcome assessment tools

Example: "The Body Image Avoidance Questionnaire (BIAQ) developed by Rosen, Srebnik, Saltzberg and Wendt (1991) was used to pretest and post-test body image disturbance. This scale is designed to measure behaviors that often accompany body image disturbance. The BIAQ contains 19-items that deal with 'avoidance of situations that provoke concern about physical appearance' (Fischer & Corcoran, 1994). Totaling the scores on each of the six point items scores the questionnaire. The possible range is 0-94. The higher the score the more avoidance behaviors are used. The internal consistency for the BIAQ is excellent, with a Cronbach's alpha of .89. It has a stable two week, test-retest reliability coefficient of .87. Further, the BIAQ has fair to good concurrent validity, with a low but significant correlation of .22 with body size estimation, and a correlation of .78 with the Body Shape Questionnaire. It also has good known-groups validity, significantly distinguishing between clinical (bulimia nervosa) and nonclinical populations and has been shown to be sensitive to changes in clients with body-image disturbance (Fischer & Corcoran, 1994)."[63]

Explanation: The instrument used to measure primary and secondary outcomes often lack evidence of reliability and validity.[64] Providing explicit details about the measurement properties of such tools will enable readers to gauge whether outcomes were measured in a sufficiently robust manner (for example, sensitive to change, valid for the condition under study), and thus the trustworthiness of findings.[65] In a review of 138 RCTs of paediatric acute diarrhoeal diseases, 87 (63%) studies explicitly stated one or more primary outcomes; none reported the use of a valid and reliable primary measure or instrument to evaluate the primary outcome.[64] Similarly, 2194 different instruments have been used in 10 000 trials in schizophrenia, of which 1142 had only been used once.[66] In another study, of non-pharmacological trials, one third of the claims of treatment superiority based on unpublished scales would not have been made if a published scale had been used.[67] In the absence of empirical evidence of the reliability and validity of outcome measures, readers should be skeptical about whether reported effect estimates reflect the intended concept and whether they are relevant or appropriate in practice.

Item 6b

Standard CONSORT item: Any changes to outcomes after the trial commenced, with reasons

CENT extension: No change

Example: No example of good reporting could be identified

Explanation: Selective outcome reporting in clinical trials has been extensively documented,[57] and similar evidence is emerging for systematic reviews.[68] In most cases, changes to planned outcomes have been shown to be associated with the nature and direction of findings, resulting in a bias of the evidence base toward favourable effect estimates—a concept termed outcome reporting bias. It is unlikely that selective outcome reporting is limited to just parallel group trials and reviews. Until the registration of N-of-1 trials is standard practice (see CENT item 23) it will remain difficult for readers to detect selective reporting in N-of-1 trial reports. Trial authors should indicate whether changes to outcomes were made (for example, added, removed, re-prioritised) and their rationale for doing so.

Sample size

Item 7a

Standard CONSORT item: How sample size was determined

CENT Extension: No change

Example: "Estimation of the needed number of cross-overs (that is, 'sample size') was based on having at least 80% power (β = 0.20) to detect a 50% reduction in symptoms, with significance testing at the α = 0.05 level. Variability in the Conners ratings was estimated based on normative data in school-aged children, which show standard deviation (SD) of 5.2 (Werry et al., 1975). Based on the baseline Conners scores of 17 and 15 in the two children, a 50% reduction in symptoms could be detected with three cross-overs, under the given model parameters. We felt that the number of cross-overs should ideally be higher, to allow for the possibility of higher intra-individual variability, and selected a target of five cross-overs."[69]

Example: "For a conventional RCT, the sample size required to detect a difference in effect of 8 on the FACIT-F fatigue subscale between MPH and placebo with 5% significance level and 80% power, using a two-sided test, is 33 per treatment group. Allowing for 30% attrition raises the sample required to 47 per group or 94 overall. Using the same information, assuming no period effect or treatment x time interaction, computer simulation of size N = 10,000 in SAS (SAS Institute Inc., Cary, NC, USA) was used to model the required sample size for the equivalent aggregated n-of-1 design. If 60% of recruited patients complete the first cycle, 50% complete the first two cycles, and 45% complete all three cycles, then 21 patients would be needed to satisfy the same significance and power requirements."[70]

Explanation:

Sample size is a distinct concept in N-of-1 trials. Within an individual trial, sample size may refer to a calculation around the number of periods comprising an individual trial or the number of measurements within a treatment period, if done. Within a series of trials, sample size may also refer to a calculation of the number of individual trials comprising the series, as indicated by the two above examples.

In a series of trials, the number of repetitions of periods across individual N-of-1 trials, and repeated sampling within periods, may be more important than the number of individual trials carried out in the series.[70] As such, whether a sample size was determined for any of these should be reported. For individual trials, the method for determining the planned number of measurements within each period or the number of periods within a trial should be reported. It is unlikely that a valid sample size for the number of periods can be calculated for trials with less than three crossovers, since the degrees of freedom for such a test would be quite small. In these instances, investigators should report the posterior probability or odds that one treatment is better than the other.

For series of N-of-1 trials, investigators who calculated a sample size should report how the intended sample size was calculated in the same manner as is done and reported in parallel group RCTs. The rationale and source of variables used to compute a sample size should be stated. For instance, if the minimum detectable difference between two treatments was obtained from group trial data, this should be stated. In addition, Type I error (that is, the probability of rejecting the null hypothesis when it is true), and power (that is, the probability of rejecting the null hypothesis when it is false) as well as whether the test is one sided or two sided should be stated.

Item 7b

Standard CONSORT item: When applicable, explanation of any interim analyses and stopping guidelines

CENT extension: No change

Example: "Patients had the option of prematurely terminating a study period if they believed they were receiving placebo and wanted to switch to the next treatment period."[32]

Example: "We did not plan to stop the evaluation early."[71]

Explanation:

Early stopping in N-of-1 trials refers to intentional stopping within a trial period based on a priori stopping rules. Such rules may be planned in anticipation for minor adverse effects or potential ineffectiveness of an intervention. Early stopping is distinct from participant withdrawal (or dropping out) from the trial completely as well as from investigator-determined exclusion. Withdrawal implies unplanned stopping of a trial, which, as in parallel group trials, may be due to a number of unanticipated reasons (such as serious adverse events).

When early stopping of a period within an N-of-1 trial occurs, the participant remains enrolled and continues on to subsequent periods. Potential reasons for early stopping include adverse effects or perceived lack of efficacy. In such circumstances, a participant may contact the clinician about how she or he is feeling and, after documentation, proceed to the next period in the a priori planned sequence; early stopping can be achieved without interfering with randomisation or blinding. The pre-specified reasons for stopping a period early should be reported.

Early stopping in N-of-1 trials (that is, selective early discontinuation of a single period) is different than early stopping in RCTs, which results in the discontinuation of the entire trial. Parallel group RCTs may be stopped early because of demonstrated treatment benefit or harm seen at interim analyses. Such trials are problematic because they tend to prematurely promote new drugs based on "random highs" (or lows) in treatment effect[72] and may misestimate treatment effects for the outcome precipitating the early stopping.[73 74]

Any a priori rules for interim analyses and early stopping, and how these were determined, should be stated. Consideration of how early stopping will affect the analysis of the trial or series should be described, particularly in longitudinal analyses

Randomisation: sequence generation

Item 8a

Standard CONSORT item: Methods used to generate the random allocation sequence

CENT extension: Whether the order of treatment periods was randomised, with rationale, and method used to generate allocation sequence

Example (randomised): "The order of medication periods was randomly assigned within each of 3 pairs of periods, according to a computer-generated randomisation schedule."[75]

Example (non-randomised): "The specific treatment approach used was alternated each week so that two sessions of one treatment were followed by two sessions of the other during the ensuing week. The two-session alternating approach was utilized to ensure that neither treatment was given an advantage in terms of the number of days between treatment and mastery probes administered in the following session."[76]

Explanation:

In traditional, parallel group RCTs, randomisation (that is, the chance based process of assigning participants to treatment or control condition) is used to ensure the even distribution of participant characteristics between groups. In N-of-1 trials, as in crossover group RCTs, the potential for confounding by covariates is eliminated due to the nature of the design—participants act as their own control. However, the assignment of treatment periods is still an important methodological consideration in N-of-1 trials, which may be done with or without employing randomisation.

In N-of-1 trials, randomisation for a patient refers to the random assignment of a treatment period to a specific treatment within a pair or block of periods of a pre-specified size. It is typically used to ensure that each treatment has an equal chance of being administered and so that patients or their health professionals cannot predict the next treatment (that is, to preserve blinding). This works well when many randomised treatment blocks are carried out or when there are a large number of patients in a series. Randomisation may be used to select the starting treatment, as done in group RCTs.

However, relying solely on randomisation is sometimes problematic, such as when an outcome unknowingly, progressively deteriorates or improves. As described by the N-of-1 panel of the Agency for Healthcare Research and Quality's DEcIDE (Developing Evidence to Inform Decisions about Effectiveness) group, an alternative to randomisation is the use of counterbalancing to select the order of treatments.[77] Many N-of-1 trials employ a standard, two treatment design (that is, treatment pair), which, when randomised, consistently places either A or B in the latter position 50% of the time (that is, ABAB or BABA for a 4 period sequence with block size of 2). If there is progressive deterioration (or improvement), the later treatment will on average result in worse (or better) outcomes than the treatment preceding it. While an N-of-1 design is recommended to be used only for stable conditions rather than progressive conditions (see box 3), it may sometimes be the only feasible mechanism of evaluation when an informed treatment decision is needed, or may be used to evaluate conditions or outcomes that are not necessarily known to be progressive. In these situations, a researcher may choose to make use of a balanced-counterbalanced design in which the sequence is selected such that treatment order (such as AB or BA) is systematically alternated (that is, ABBA or BAAB) so that neither treatment suffers a worse fate than the other, solely based on the ordering within a treatment block. Whether randomisation, counterbalancing, or another mechanism was used to determine the order of treatments should be described by authors, along with a rationale and how it was achieved.

As in group RCTs, the method used to decide on the treatment sequence is an essential method in N-of-1 trials. Whether or not randomisation was used should be reported along with the rationale for the selected treatment order with special attention for how the effect of time (that is, period effect) was addressed. The mechanism used to generate the randomisation sequence, such as computer-generated, random numbers table, coin toss or other random selection process, should be described in enough detail to enable readers to gauge whether the method used was robust. Simply stating that randomisation was used is not sufficient. Additionally, authors should describe other details of randomisation, as they apply (see CENT item 8b).

Item 8b

Standard CONSORT item: Type of randomisation; details of any restriction (such as blocking and block size)

CENT extension: When applicable, type of randomisation; details of any restrictions (such as pairs, blocking)

Example: "Following a 2-week run-in period, quinine sulphate and matched placebo capsules were compared in three 4-week treatment blocks (each block consisting of 2 weeks active drug and 2 weeks placebo in random order)."[78]

Explanation: If randomisation was used to generate treatment sequence, the unit of randomisation, such as within a pair or block, or if treatments were simply alternated after randomly assigning the starting treatment, should be reported.[79] If blocking was used, the block size should be reported as well as whether the size was fixed or randomly decided. As with parallel group RCTs, if the trialists became aware of the block size(s) during the trial, that information should be reported as such knowledge could lead to code breaking.[27] Whether a predetermined ratio other than 1:1 was used should also be reported, along with a rationale for the type of randomisation used and any associated limitations.

Item 8c

Standard CONSORT item: None (new for CENT)

CENT extension: Full intended sequence of periods

Example: "Within each pair, the sequence was randomized by the pharmacist who had no contact with the patient. The actual sequence was [drug, placebo], [drug, placebo], [placebo, drug]...."[80]

Explanation:

In existing N-of-1 trial reports, it is common practice to state the sequence completed in the trial in the results section. However, the generation of a treatment sequence is a method carried out before the start of the trial and so should be reported in the methods section of a trial report, even if it changed after the start of the trial. The completed sequence and reasons for changes, if any, should be reported as per CENT item 14a (numbers and periods analysed). This guidance is in line with CONSORT and other reporting guidance.

For series of N-of-1 trials, where the sequence is different for each individual trial, it may not be possible to report the planned sequence for each trial in the text. Sequences for each individual trial may instead be included as an appendix.

Randomisation: Allocation concealment mechanism

Item 9

Standard CONSORT item: Mechanism used to implement the allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned

CENT extension: No change

Example: "Active and placebo medication was issued in 'Webster packs' manufactured by Webstercare (packed and sealed medication according to individual's dosage requirements, such as by time of day, day of the week and week number) by a local pharmacist not participating in the trials in accord with the randomisation schedule supplied by the database manager. Patients, participating GPs and research staff were blinded to all randomisation and packaging procedures until completion of each trial."[43]

Explanation: See CONSORT 2010 item 9

Randomisation: Sequence implementation

Item 10

Standard CONSORT item: Who generated allocation sequence, who enrolled participant, and who assigned participant to interventions

CENT extension: No change

Example: "The randomisation code for eformoterol and placebo turbuhalers was independently supplied in opaque envelopes, with allocation on study entry in order to blind subjects, research staff, and two respiratory physicians who inspected the outcome data for each individual participant."[39]

Explanation: See CONSORT 2010 item 10.

Blinding

Item 11a

Standard CONSORT item: If done, who was blinded after assignment to interventions (for example, participant, care providers, those assessing outcomes) and how

CENT extension: No change

Example: "Both the patients and the researcher interacting with them and conducting the analyses were blinded to when patients were taking the active drug or the placebo."[78]

Explanation: See CONSORT 2010 item 11a.

Item 11b

Standard CONSORT item: If relevant, description of the similarity of interventions

CENT extension: No change

Example: "Placebos were identical in appearance, texture and weight to the corresponding active medication and contained 3% active valerian to ensure identical odour."[43]

Explanation: See CONSORT 2010 item 11b.

Statistical methods

Item 12a

Standard CONSORT Item: Statistical methods used to compare groups for primary and secondary outcomes

CENT extension: Methods used to compare data between interventions for primary and secondary outcomes

Example: "An N-of-1 RCT was considered positive if the CRQ dyspnea score was higher (i.e., less dyspnea) during the oxygen treatment period in all three pairs and if the difference between oxygen and placebo periods was 0.5 or greater during at least two of the three pairs. Analysis of each N-of-1 RCT included a paired *t* test…. Mean oxygen and placebo gas usage was determined by averaging the amount used over each of the periods for each patient."[32]

Explanation:

In line with recommendations made by the International Committee for Medical Journal Editors (ICMJE) and the CONSORT group, analytical methods should be described "with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results."[81] Two broad analytic approaches are used in N-of-1 trials: visual assessment and statistical analysis. Many N-of-1 trials authors provide a visual representation of the data, allowing readers to inspect the slope, variability, and patterns of the data, potential treatment overlap between periods, and the overall reliability and consistency of treatment effects.[82 83] Analytical aids such as a line of best fit may sometimes be used to facilitate interpretation of visually presented data. If done, authors should describe how the analysis was carried out.

With the increasing focus on evidence based medicine, some argue that the statistical determination of effect sizes (magnitude and direction) provides trial findings in manner that can be used more easily understood and used in clinical decision making and by future researchers (that is, for meta-analysis).[83] It is not usual for visual assessment to precede and inform the need for statistical analysis, based on an implicit rule (that is, if one treatment seems to yield better outcomes than another most of the time). Another rule may be based on the number of data points, since, as more data are displayed on a graph (such as multiple measurements per period), visual assessment becomes more difficult unless there is very little within-sample variation.

In N-of-1 trials special consideration of the handling and summarising of data is required since data are typically presented for each outcome, in addition to estimation of treatment effect. Authors may or may not choose to statistically summarise individual trial data. If done, several options are available and authors should state which data were summarised and how this was done. For instance, authors may plan to summarise multiple measurements of a given outcome within a period (that is, length of time when only one treatment is given) or measurements from all periods of a recurring treatment overall (that is, combined data from all periods of a given treatment). If data are combined, methods for summarising (such as the mean) and estimating their variance (such as standard deviation) should be reported. Whether a visual or statistical approach to analysis (or both) was used (and a description of how each method was carried out, as specified above) should be reported. If rules were used to determine the need for statistical analysis, authors are encouraged to report what the rule(s) is and whether it was specified before seeing the data (a priori).

A number of statistical approaches for determining treatment effect sizes have been documented in N-of-1 trials with little consensus.[84-86] Authors should describe the selected measure used to generate effect estimates, that is, to compare data between treatments (such as mean difference), and which measure will be used to indicate the precision (uncertainty) of the estimate.[32 87 88] As in group trials, a 95% confidence interval is standard, but occasionally other levels are used depending on the level of conservativeness needed. Many biomedical journals require or strongly encourage the use of confidence intervals.[81] They are especially valuable in relation to differences that do not meet conventional statistical significance, for which they often indicate that the result does not rule out an important clinical difference. The use of confidence intervals has increased markedly in recent years, although not in all medical specialties.[87] Although P values may be provided in addition to confidence intervals, results should not be reported solely as P values.[89] If authors choose to also report P values, the actual value (such as P=0.003) should be given, rather than whether it is above or below an arbitrarily chosen point (such as P<0.05).

If both continuous and dichotomous or categorical outcomes are measured, authors should distinguish their approaches for analysing each type of data. For reports combining data from a series of N-of-1 trials, refer to CENT item 12b.

If sensitivity analyses were planned and carried out, authors should state what analyses were done (that is, excluding outlying data points, periods stopped early, etc).

Item 12b

Standard CONSORT item: Methods for additional analyses, such as subgroup analyses and adjusted analyses

CENT extension for series only: If done, methods of quantitative synthesis of individual trial data, including subgroup analyses, adjusted analyses, and how heterogeneity between participants was assessed (for specific guidance on reporting syntheses of multiple trials, please consult the PRISMA Statement)

Example: "Bayesian analyses combining N-of-1 trials employed a 2-level random effects model to describe the posterior distributions of treatment effectiveness. For these analyses we assumed a common within-patient variance. More complex variance structures did not improve model performance. Analyses used both non-informative and, separately, informative priors derived from published trial results."[90]

Example: "To address the effect of oxygen on the entire group, we conducted repeated-measures analysis of variance, examining the effects of treatment, pair, and the treatment-pair interaction. Mean oxygen and placebo gas usage was determined by averaging the amount used over each of the periods for each patient. The correlation between mean oxygen and mean placebo gas usage was calculated for the entire group as an intraclass correlation coefficient (ICC)."[32]

Explanation: This item is only applicable for reports of series of N-of-1 trials, where authors carried out a quantitative synthesis of data from more than one trial. Doing so may provide readers with both an overall average effect, potentially stratified by participant or study characteristics, and revised estimates of each individual's outcomes that are informed by the results of other participants.[8] If subgroup analyses were performed, for instance, stratification by participant or study characteristics, this should be explicitly stated.

There is not yet a single, widely recognised approach to synthesising data from N-of-1 trials, but it should be noted that methods are likely distinct from the synthesis of group trial data. Meta-analytic methods for combining individual trials have been explored and reported by some,[8 47 91 92] but further exploration is needed. Authors should report the approach used (such as Bayesian or frequentist) and the models used (such as fixed or random effects). In addition, the summary measures used and the level at which the resulting effect estimate would be considered significant (such as 95% confidence interval) should be reported. Since the number of observations taken in each N-of-1 study is relatively small, and results may vary substantially between individuals, the explicit statistical method used to explore relative heterogeneity both within and between individuals should be reported.[47] In some instances, subgroup and sensitivity analyses may be used to explain differences between treatments among individuals in a series.

Authors should be clear about whether raw data or summary data from individual trials were used to determine group estimates. In line with individual patient data meta-analyses,[93] raw data is preferable. Specific reporting guidance for systematic reviews (and retrospective meta-analyses) that include N-of-1 trials and series is planned. However, in the interim, authors are encouraged to refer to the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analysis) Statement,[94 95] as the overall review process and approach to reporting it will be similar.

Item 12c

Standard CONSORT item: None (new for CENT)

CENT extension: Statistical methods used to account for carryover effect, period effects, and intra-subject correlation

Example: "the averages of these differences for the subjects who began the alternating-treatment sequence with amphetamine and for those who began with caffeine were compared in a two-sample *t*-test crossover analysis as described by Hills and Armitage and Armitage and Berry. This method accounts for sequence effects in addition to drug effects."[96]

Example: "Carryover and time trend significance were tested using random-effects regression models that included treatment pattern variables and time together with a time-by-treatment interaction term, respectively"[90]

Explanation:

A common issue for N-of-1 trials (a form of crossover study) is whether the outcome in the current period is influenced by the treatment given in the previous period. If the effects of an intervention are long lasting, they may carry over into the next period if the washout period for a given treatment is insufficient or absent. This has the potential to result in biased estimates of differences between interventions. Authors should report whether a carryover effect was explored and how it was accounted for in the analysis. A period effect is a change that would have occurred even in the absence of treatment, due to time; if examined, methods of doing so should be reported.[97] See box 4 for definitions of period effect, carryover effect, and intra-subject correlation.

Authors should also be aware that multiple observations in a single patient are not independent and that treating them as "independent" data is a serious problem which may result in increased type I or type II error,[98] as in group crossover RCTs.[99] The method used to determine whether data are autocorrelated (as in time series analysis)[100] as well as the method(s) used to account for it (such as long baseline and post-intervention data)[101] should be described.

Results

Participant flow (a flow diagram is strongly recommended)

Item 13a

Standard CONSORT item: For each group, the numbers of participants who were randomly

assigned, received intended treatment, and were analysed for the primary outcome

CENT extension: Replaced by item 13a.1 and 13a.2

Item 13a.1

CENT extension: Number and sequence of periods completed and any changes from original plan, with reasons

Example: "The completed treatment order therefore was: placebo, valproic acid, placebo, placebo, valproic acid, placebo, missed, valproic acid."[102]

Example: "The order in which treatments (d=drug, p=placebo) were given was: p d d p d p p d p d."[56]

Note: No example reporting changes from the original plan could be identified.

Explanation: The number and order of treatments as received by a patient may differ from what was determined by randomisation or otherwise planned a priori. The actual completed sequence may be represented in a trial pictorial (see CENT item 17a.1), but it is important that authors distinguish whether changes were made to the original plan, and why, so readers can judge whether changes are justified or may be associated with bias, if, for instance, the change was associated with a treatment effect.

Item 13a.2

CENT extension for series only: The number of participants who were enrolled, assigned to interventions, and analysed for the primary outcome

Example: Flow diagram (fig 5⇓) from Nonoyama et al.[32]

Explanation:

The phases for a series of N-of-1 trial are akin to those of a parallel group RCT (such as recruitment, enrollment, treatment allocation, follow-up, analysis). The design and conduct of some N-of-1 trial series is straightforward, and the flow of participants, particularly when there are no exclusions or losses to follow-up (item 14c), through each phase of the study can be described adequately in a few sentences. A diagram depicting the flow of participants through phases of an N-of-1 trial series (fig 4⇓, CENT flow diagram) will provide a snapshot of when and why participants were excluded, and whether such exclusions are indicative of potential bias in treatment estimates, and will assist in gauging generalisability of findings.

Participants who were excluded or dropped out of a trial because of a treatment related harm relating to their particular condition are not necessarily representative of the broader patient population. However, the number of people assessed for eligibility but excluded (and reasons why) is a useful indicator of whether inclusion criteria, and thus subsequent findings, are appropriate or applicable to a real life scenario.

Item 13b

Standard CONSORT item: For each group, losses and exclusions after randomisation, together with reasons

CENT extension for series only: Losses or exclusions of participants after treatment

> **Box 4: Considerations of crossover designs**
>
> *Period effect*—A change that would have occurred even in the absence of treatment, due to time[97]
>
> *Carryover effect*—The persistence, into a later period of treatment, of some of the effects of a treatment applied in an earlier period[121]
>
> *Intra-subject correlation*—The variation, exhibited by a single person, on repeated measurements. As an illustrative example, observed resting pulse rate will show considerable minute-to-minute variation but little evidence of any trend over time. Successive within-subject values are unlikely to be independent[122]

assignment, with reasons, and period in which this occurred, if applicable

Example: "Complete results for 7 of the 8 study weeks were obtained, because the patient was ill and away from school on the Friday of the 7th week."[102]

Example: "One patient withdrew halfway through the final period-pair due to unrelated personal reasons. A second patient was hospitalized for gastrointestinal bleeding during the second period-pair, and warfarin treatment was discontinued.… For both patients, data collected during their participation in the study were included in the analysis."[33]

Example: "The study was terminated on day 5 of the last placebo period because of epistaxis."[103]

Explanation:

With respect to individual N-of-1 trials and series of trials, exclusions and losses to follow-up may be incurred at two levels. In a series of N-of-1 trials, a particular participant may be excluded or lost to follow-up after treatment allocation (see item 13a.2). In the case of an individual trial, exclusions or losses to follow up may occur when a participant fails to complete a particular period or set of periods yet continues on to complete the remainder of the trial as planned. In the latter case, investigators may choose to analyse completed pairs or blocks of treatment. If this is done, the periods excluded from analysis should be reported.

In N-of-1 series, when data from participants who were lost to follow-up are excluded from the analysis entirely, erroneous conclusions can be reached. Knowing the number of participants who did not complete the trial and whose data are not included in the analysis permits the reader to assess to what extent the estimated efficacy of therapy might be under or overestimated in comparison with ideal circumstances. Intention-to-treat analysis is fundamental to preserving the pre-planned statistical power of research studies, any reduction in which may cause treatment effects to go undetected or be wrongly estimated.

Authors should also distinguish between exclusion of participants based on predetermined or investigator-determined criteria (such as ineligibility, withdrawal from treatment, and poor adherence to the trial protocol) and attrition resulting from loss to follow-up, which is often unavoidable.

Recruitment
Item 14a

Standard CONSORT item: Dates defining the periods of recruitment and follow-up

CENT extension: No change

Example: "Between April and November 2001, [physicians] were asked to select patients from their medical records who met the following criteria…. The study comprised of a series of N-of-1 trials with a duration of ten weeks."[29]

Explanation: In an individual N-of-1 trial, purposeful "recruitment" generally does not happen; rather the decision to conduct an N-of-1 trial may stem from the need to answer a clinical question based on uncertainty of treatment effect for a particular individual. The concept of recruitment is more applicable to series of N-of-1 trials led by a broader, population-based question for which RCT evidence is lacking or inapplicable. The period of recruitment for a series of N-of-1 trials should be reported as well as whether and how long a follow-up period lasted. The start and end dates of the trial should be reported, although authors should be mindful of the potential for patient identification when combined with other personal information (items 4a, 4b, and 15).

Item 14b

Standard CONSORT item: Why the trial ended or was stopped early

CENT extension: Whether any periods were stopped early, and whether the trial was stopped early, with reason(s)

Example: No example of good reporting was identified

Explanation:

N-of-1 trials, may, like parallel group trials, end after their pre-specified duration or end earlier than planned for a number of different reasons. The reporting of this item overlaps, in parts, with item 14c. Specifically, losses and exclusions during an individual N-of-1 trial essentially accounts for some of the reasons that early stopping may occur. Data may not always be lost and, when available, may be very informative. For example, a participant may consistently stop a period early in a blinded fashion due to deterioration or adverse effect. Knowledge of such information will allow readers to judge whether the early stopping of a period is important and has any impact on the reported findings.

In other instances, reasons for early stopping may be planned or anticipated (item 7b) and may or may not be related to trial results. Interim analyses may reveal a benefit or no difference between treatment alternatives (that is, futility) leading investigators to decide to end a trial early. Timing of interim analysis, when the trial was stopped (indicating the specific period), and reasons why should be stated.

For the same reasons as in parallel group trials,[73] interim analyses should prudently employed, interpreted with caution and explicitly reported as such. In N-of-1 trials, interim analyses may reduce the number of repetitions of treatment periods (that is, sample size), overestimate treatment effect, or uncover spurious associations.

Baseline data
Item 15

Standard CONSORT item: A table showing baseline demographic and clinical characteristics for each group

CENT extension: No change (no table necessary for individual trials)

Example: Table of baseline characteristics (fig 6⇓) from Smith et al.[39]

Explanation: Whether an individual or series of trials was done, it is important to provide the reader with underlying demographic and clinical information for each participant to aid in interpreting appropriate application of the results. When reporting a series of N-of-1 trials, it may not be feasible to present individual characteristics for each patient, but a table with descriptive summary data for baseline characteristics should be included. For small series (n≤10), authors should aim to provide characteristics for each patient. The items of most interest include demographics, clinical history, socioeconomic data, prior medications taken, baseline health measurements, and any other baseline characteristic specific to the trial at hand that authors think may aid readers to gauge the generalisability of the findings.

Numbers analysed
Item 16

Standard CONSORT item: For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups

CENT extension: For each intervention, number of periods analysed

Example: No example of good reporting could be identified.

In addition for series: If quantitative synthesis was performed, number of trials included

Example: "For both patients, data collected during their participation in the study were included in the analysis."[30]

Explanation:

Analysis according to the trial as planned (such as intention-to-treat including data from all measurements or periods) rather than as actually occurred (such as per protocol) has become standard practice in group RCTs. Authors should report both the number of periods from which data were analysed as well as how missing periods (or periods during which outcomes were unavailable) were analysed. Furthermore, as indicated in item 3a, authors should state the planned number of periods to be included in analysis, for each intervention, so readers can compare planned versus actual and make their own judgments. Authors should also state whether information for all outcomes is available for all periods, or specify which periods were included in analysis for each outcome.

Intention-to-treat analysis allows data from all participants to factor into the analysis, regardless of the reason for not following the protocol. However, since intention-to-treat analysis can provide very conservative estimates if data are missing, authors should explicitly describe and rationalise any methods of imputation used.

For series of trials, if quantitative synthesis (such as meta-analysis) is performed, the number of trials included in the synthesis should be reported.

Outcomes and estimation
Item 17a

Standard CONSORT item: For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)

CENT extension: Replaced by items 17a.1 and 17a.2

Item 17a.1

CENT extension: For each primary and secondary outcome, results for each period; an accompanying figure displaying the trial data is recommended

Example: Results (fig 7⇓) from Estrada et al.[104]

Example: Results (fig 8⇓) from Nonoyama et al.[32]

Explanation: Outcome data displayed in a figure or table are often clearer than descriptions in text. A clear description of the results for each treatment in each pair or block should be provided for each primary and secondary outcome. For each outcome, if repeated measures within a period are combined, a summary measure (such as mean or median) and indication of variance (such as standard deviation or interquartile range) should be reported.

One possible way of presenting individual patient data is to plot all outcome measurements for each period over the trial duration, as shown in the example from Estrada et al (fig 7⇓) above and figure 3⇓ (N-of-1 trial pictorial). Outcome measurements are plotted on the Y-axis, units of time (days, weeks, etc) along the X-axis, and vertical lines or some other distinction made between treatment periods and pairs or treatment blocks. For a series of N-of-1 trials, authors may wish to present individual trial pictorials in an appendix. For individual and series of N-of-1 trials, a table reporting each person's complete raw data is strongly recommended (as an appendix if necessary), consistent with the current open data movements.[105 106]

Item 17a.2

CENT extension: For each primary and secondary outcome, the estimated effect size and its precision (such as 95% confidence interval)

Example: Individual trial data (fig 9⇓) from Avins et al.[107]

Example: Individual trial data in a series (fig 10⇓) from Zucker et al.[90]

In addition for series: If quantitative synthesis was performed, group estimates of effect and precision for each primary and secondary outcome

Example: Series data (fig 11⇓) from Nonoyama et al.[32]

Example: Series data (fig 12⇓) from Coxeter et al.[43]

Explanation:

When reporting on the treatment effect size, comparison between a summary of all measures for each treatment (that is, within and across periods) is typical. Effect size estimates should be accompanied by a confidence interval. For binary outcomes, the effect size may be represented by a relative risk (risk ratio), odds ratio, or risk difference; a mean difference is typically calculated for continuous data. Authors should provide a confidence interval to indicate the precision (uncertainty) of the estimate.[87] If P values are reported, the actual value (such as P=0.003) rather than whether it is above or below an arbitrarily chosen point (such as P<0.05) should be provided.

For an individual trial, authors should report the effect size and confidence interval for each outcome (fig 9⇓). For series of

trials, authors may wish to report individual trial effect sizes separately (fig 10⇓) or pool data for each outcome among series participants (figs 11⇓ and 12⇓). More information on the methods for combining N-of-1 trial data can be found in the Agency for Healthcare Research's user's guide for N-of-1 trial design and implementation.[77] When effect sizes are combined between trials, authors should still report calculated effect sizes and precision (such as 95% confidence interval).

Effect sizes should be reported for all planned primary and secondary end points, and for all participants if a series was carried out, not just statistically significant or interesting effects. The selective reporting of outcomes within population based RCTs is a widespread and serious problem (see CENT item 6a.1). As previously stated, although we are unaware of empirical evidence of selective reporting of data for only statistically significant or interesting outcomes in N-of-1 trial reports, it is conceivable and authors should avoid this practice.

Item 17b

Standard CONSORT item: For binary outcomes, presentation of both absolute and relative effect sizes is recommended

CENT extension: No change

Example: No example of good reporting identified.

Explanation: As in group crossover trials, the measurement of binary outcomes in N-of-1 trials is rare as it has the potential to be problematic. If a binary outcome is measured only once within a period, there is a higher potential for residual effects to influence the next treatment, unless sufficient washout can be guaranteed or carryover accounted for in the analysis.[108] Readers should be mindful of this, and authors should report measurement frequency, as indicated in CENT item 6a.2. If binary outcomes are measured and combined within or between (that is, in series) trials, authors should report absolute and relative effect sizes and whether and how carryover effect was assessed or addressed.

Ancillary analysis
Item 18

Standard CONSORT item: Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory

CENT extension: Results of any other analyses performed, including assessment of carryover effects, period effects, and intra-subject correlation

Example: "Although there was an improvement over time for the CRQ dyspnea (p = 0.05) and the CRQ mastery (p = 0.001), these effects were unrelated to the gas mixture being used, with no main effects of oxygen, nor any interaction between treatment and pair for any of the CRQ domains. Furthermore, the upper boundary of the CI excluded a mean difference greater than the MID for all four CRQ domains (Table 2, Figure 3). There were no significant differences between oxygen and placebo, or improvements with time for any domains of the SGRQ."[32]

Example: "The tests for carry-over effect disclosed no significant differences in response measures in general or in any of the subgroups."[109]

In addition for series: If done, results of subgroup or sensitivity analyses

Example: "We tested for interactions of treatment effect and enrollment site (center vs. community) or any of the selected patient characteristics. No statistically significant associations were found in our study population (data not shown). Additionally, no significant interactions between time and treatment effect or between treatment order and treatment effect were identified using random-effects regression models (data not shown)."[90]

Note: No example could be identified that reported on a subgroup or sensitivity analysis.

Explanation: In addition to analyses of treatment effects, additional analyses may have been carried out, whether stated a priori or not, to help authors further explore their data. All such analyses, as well as whether they were planned, should be reported. Additional analyses may include a test for presence of carryover effect or period effect (see CENT item 12c) similar to those that might be carried out for a crossover trial.[97]

Where data from series of N-of-1 trials are quantitatively synthesised, subgroup and sensitivity analyses as well as meta-regression techniques may be used to explain heterogeneity and interpret the data and should be reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines.[94 95]

Harms
Item 19

Standard CONSORT item: All harms or unintended effects in each group (for specific guidance see CONSORT for harms)

CENT extensions: All harms or unintended effects for each intervention (for specific guidance see CONSORT for harms)

Example: "Only one adverse event-severe foot/ankle swelling on celecoxib-resulted in withdrawal. Nine patients reported more adverse events while on SR paracetamol than on celecoxib, and five reported more while on celecoxib than on SR paracetamol. In the other 25 patients, there was no difference in the prevalence of adverse events reported. The most common adverse events on celecoxib were headache (54%), loss of energy (54%), indigestion (36%) and constipation (32%); and on SR paracetamol were loss of energy (51%), headache (49%) and constipation and indigestion (44%) (Table 5). There were differences between the two drugs in terms of stomach pain (15% for celecoxib vs. 27% for SR paracetamol) and vomiting (2% for celebrex vs. 7% for SR paracetamol)."[110]

Explanation: N-of-1 trials are an important mechanism for detecting harms that may uniquely occur in specific patients, which may or may not have been previously detected in group trials or other epidemiological studies. In some instances, an N-of-1 trial will have been designed specifically to confirm if an adverse event is attributable to a given therapy (for example, a patient with hypertension, asthma, and a persistent cough undergoes a systematic N-of-1 evaluation to confirm if his antihypertensive medication is causing the cough). If the N-of-1 trial identifies a harm or otherwise unintended effect from the treatment administered, authors must document this essential information in their trial report. Specifically, the nature of each harm, its severity, along with the period and treatment block during which it occurred should be indicated. Equally, if no harms were detected, authors should state that. Readers need information about harms and lack of harms, in addition to

information about the benefits of interventions in order to make informed decisions. Although serious harms are generally rare, reporting harms is important even if they occur in a single patient.

Discussion

Limitations
Item 20
Standard CONSORT item: Trial limitations, addressing sources of potential bias and imprecision, and if relevant, multiplicity of analyses
CENT extension: No change

Example: "A second limitation relates to the duration of the trial; the long-term impact of the medication cannot be addressed. A further limitation relates to the low number of treatment pairs."[111]

Explanation: See CONSORT item 20.

Generalisability
Item 21
Standard CONSORT item: Generalisability (external validity, applicability) of the trial findings
CENT extension: No change

Example: "The findings of this study are exploratory in nature. Given the small sample size, these results should not be generalised beyond the sample. The findings, however, suggest the potential benefit of TT for women with migraine headaches. Participants in the study experienced the beneficial effects of decreased migraine frequency and increased relaxation levels in response to the TT intervention, with no documented adverse effects. These findings warrant further exploration of the effect of TT on more diverse migraine headache populations and with a larger sample."[112]

Explanation: See CONSORT 2010 item 21

Interpretation
Item 22
Standard CONSORT item: Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
CENT extension: No change

Example: "In other studies, 31 (72%) of 43 children had a good response to methylphenidate, and 48 (68%) of 70 children showed improvement on methylphenidate in one of two 2-week periods. Of 94 children treated with methylphenidate, 70 (74%) demonstrated a positive response. Five (50%) of 10 children with ADHD and IQs ranging from 48 to 77 responded positively to methylphenidate. Our placebo-controlled response rates are lower than these, probably because of our different selection criteria and more strict criteria for response. We also tested both methylphenidate and dexamphetamine against each other and placebo, which none of the other studies did. In our study, only 20% of the patients withdrew, which is much lower than the withdrawal rates previously reported for N-of-1 trials in Australia (40% and 37%). This may be related to our ability to shorten the trials because stimulants are eliminated so quickly from the body."[113]

Explanation: The explanation for this item does not differ from CONSORT item 22. Authors may also wish to report whether

the eventual treatment decision was consistent with findings from the N-of-1 evaluation.

Other information

Registration
Item 23
Standard CONSORT item: Registration number and name of trial registry
CENT extension: No change

Example: "The study was registered in EudraCT database under the number 2009-011736-35."[114]

Explanation:

This item is no different from CONSORT 2010 item 23, although we feel that further explanation is warranted. Much of the larger discussion around a priori study registration has focused on registration of group RCTs. The registration of RCT protocols has substantially improved over the past decade since the introduction of a mandatory registration policy by ICMJE journals publishing trials[115] and the US Food and Drug Administration.[116] Trial registration was essential to the initial characterisation of selective reporting of outcome in published RCTs, now known to be a widespread problem.[57] Similarly, registration of N-of-1 trials will allows readers to compare a priori methods documented in registries against those in a final report (published or not), if desired, to determine if changes were made and whether they affect reported findings.

Based on the pattern of early clinical trial registration, it is thought that only a small proportion of clinically oriented N-of-1 trials are registered and subsequently submitted for publication. Publication may or may not be related to the favourability of the outcome of the trial, risking the potential for publication bias.[79] Ideally, all N-of-1 trials should be registered so readers can be aware that they were conducted, even if not published. Ideally all N-of-1 trials should be published so that they can become part of the scientific record regardless of the outcome of the trial.

Registration of N-of-1 trials may also inform the improvement of methods and the development of study protocols for N-of-1 trials. The availability of N-of-1 trial protocols is not yet commonplace, but existing databases can accommodate the registration of N-of-1 trial protocols. A recent systematic review indicates 97% of published trials do not state whether a trial is registered, and only one trial was identified in which registration details were provided in the report (see example above).[12] Authors are encouraged to register their N-of-1 trials before conducting them, and including registration information (registry name and unique identifier) in manuscripts submitted for journal publication or public consumption.

Protocol
Item 24
Standard CONSORT item: Where the full trial protocol can be accessed, if available
CENT extension: No change

Example: "The protocol for this N-of-1 trial and supporting CONSORT checklist are available as supporting information; see Protocol S1 and Checklist S1."[117]

Explanation: This item is no different from CONSORT 2010 item 24, but we have provided additional rationale here. Trial protocols often contain more detailed information than their accompanying registry entries. Increasingly, the comparison of

trial protocols (or registry entries) with completed reports shows a plethora of changes to trial design and outcomes, not always made clear by authors in the final report, many of which represent a form of selective reporting bias.[57] For N-of-1 trials, the extent of this problem is yet unknown. Authors are asked to indicate where a protocol for the reported trial can be found.

Funding
Item 25
Standard CONSORT item: Sources of funding and other support (such as supply of drugs), role of funders
CENT extension: No change

Example: "Funding: this study was supported unconditionally by Leo Pharma, The Netherlands."[29]

Explanation: See CONSORT 2010 item 25.

Discussion

The CENT guidelines will allow decision-makers to make better use of reported N-of-1 data. A standardised approach to reporting of N-of-1 trials will facilitate increased clarity in the communication of N-of-1 trial methods, analysis, and outcomes. This will provide readers, such as clinicians, with enough information to judge the methodological rigour of the trial, whether treatment outcomes may have been affected by bias, and, ultimately, whether to employ an intervention in clinical practice.

N-of-1 trials have the potential to provide treatment information for patients who are typically excluded from evaluation in large scale RCTs, but good reporting is crucial.

The CENT guidance will also improve the usability of N-of-1 trial data by researchers, such as systematic reviewers, whose work in turn helps policymakers make decisions about healthcare. Usability of primary evidence in systematic reviews has been a longstanding, recognised problem.[118-120] If well reported, N-of-1 trials are ideal candidates for inclusion in systematic reviews since they are a source of methodologically rigorous data on treatment evaluation for patients who are often omitted from RCT evaluation; doing so may broaden the applicability of systematic reviews. Furthermore, while the primary aim of N-of-1 trials is not necessarily to provide generalisable information, if they are more broadly representative and are combined appropriately, they may be useful in making decisions about healthcare policy where RCT data do not apply or exist.

CENT will allow other researchers and decision makers to make better use of N-of-1 data. For instance, those responsible for state or provincial or hospital formulary decision making can make better use of N-of-1 trials to help determine which therapies should be eligible for coverage for specific individuals, rather than the "all or none" approach that is often used. When new uses for existing drugs emerge, N-of-1 trials may be a quick and less expensive mechanism with which to explore whether off label drug use might be effective (that is, prior to or instead of a large scale RCT). The availability of such evidence, if reported well, may also be of use to regulators when making decisions about additional conditions of use for particular treatments.

The CENT guidelines may also have the potential to affect the way that N-of-1 trials are designed before their reporting becomes a consideration, if they are consulted by researchers earlier in the research process. Furthermore, if researchers adhere to newly developed standards of N-of-1 design and implementation,[11] which are in line with CENT, compliance with CENT will be made easier.

Optimising the reporting of N-of-1 trials such that information can be accurately and transparently gleaned from them will increase their clinical value as an evaluation tool for making evidence-based decisions and promoting evidence-based practice.

1 Gitterman DP, Hay WW Jr. That sinking feeling, again? the state of national institutes of health pediatric research funding, fiscal year 1992-2010. *Pediatr Res* 2008;64:462.

2 Gagne JJ, Thompson L, O'Keefe K, Kesselheim AS. Innovative research methods for studying treatments for rare diseases: methodological review. *BMJ* 2014;349:g6802.

3 Fortin M, Dionne J, Pinho G, Gignac J, Almirall J, Lapointe L. Randomized controlled trials: do they have external validity for patients with multiple comorbidities? *Ann Fam Med* 2006;4:104-8.

4 Zarin DA, Young JL, West JC. Challenges to evidence-based medicine. *Soc Psychiatry Psychiatr Epidemiol* 2005;40:27-35.

5 Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol* 2009;62:499-505.

6 Fortin M, Bravo G, Hudon C, Vanasse A, Lapointe L. Prevalence of multimorbidity among adults seen in family practice. *Ann Fam Med* 2005;3:223-8.

7 Van den Akker M, Buntinx F, Metsemakers JFM, Roos S, Knottnerus JA. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J Clin Epidemiol* 1998;51:367-75.

8 Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, Lau J. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *J Clin Epidemiol* 1997;50:401-10.

9 Guyatt G, Jaeschke R, McGinn T. Part 2B1: therapy and validity. N-of-1 randomized controlled trials. In: Guyatt G, Rennie D, Meade MO, Cook DJ, eds. *Users' guides to the medical literature* . American Medical Association, 2002:275-90.

10 OCEBM Levels of Evidence Working Group. *The Oxford levels of evidence 2* . Oxford Centre for Evidence-Based Medicine, 2011. www.cebm.net/index.aspx?o=5653.

11 Kravitz RL, Duan N, eds, and the DEcIDE Methods Center N-of-1 Guidance Panel. *Design and implementation of N-of-1 trials: a user's guide* . AHRQ Publication No 13(14)-EHC122-EF. Agency for Healthcare Research and Quality, 2014. www. effectivehealthcare.ahrq.gov/N-1-Trials.cfm.

12 Punja S, Bukutu C, Shamseer L, Sampson M, Hartling L, Urichuk L, et al. Systematic review of the methods, statistical analysis and meta-analysis of N-of-1 trials [dissertation]. University of Alberta, 2014.

13 Hersen. *Single case experimental designs: strategies for studying behaviour change* . Peragmon Press, 1976.

14 Kazdin A. *Single case research designs: methods for clinical and applied settings* . Oxford University Press, 1982.

15 Barlow DH, Hersen M. *Single case experimental designs: strategies for studying behavior change* . Pergamon Press, 1984.

16 Tate RL, McDonald S, Perdices M, Togher L, Schultz R, Savage S. Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the single-case experimental design (SCED) scale. *Neuropsychol Rehabil* 2008;18:385-401.

17 Tate RL, Perdices M, Rosenkoetter U, Wakim D, Godbee K, Togher L, et al. Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: the 15-item risk of bias in N-of-1 trials (RoBiNT) scale. *Neuropsychol Rehabil* 2013;23:619-38.

18 Guyatt G, Jaeschke R, McGinn T. N-of-1 randomized controlled trials. In: Guyatt G, Rennie D, Meade MO, Cook DJ, eds. *Users' guides to the medical literature: a manual for evidence-based clinical practice* . 2nd ed. McGraw-Hill, 2008.

19 Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Med Care* 2011;49:761-8.

20 Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267-76.

21 Tate R, Togher L, Perdices M, McDonald S, Rosenkoetter U, on behalf of the SCRIBE Steering Committee. Developing reporting guidelines for single-case experimental designs: the SCRIBE project. *Brain Impairment* 2013;13:135.

22 Schulz K, Altman D, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;8:18.

23 Turner L, Moher D, Shamseer L, Weeks L, Peters J, Plint A, et al. The influence of CONSORT on the quality of RCTs: an updated review. 9th Annual Cochrane Canada Symposium, February 16-17, 2011.

24 Vohra S, Shamseer L, Sampson M, Bukutu C, Schmid CH, Tate R, et al. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ* 2015;350:h1738.

25 Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med.* 2010;7:e1000217.

26 Enhancing reproducibility. *Nature Methods* 2013;10:367.

27 Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010;63:e1-37.

28 Price JD, Grimley EJ. An N-of-1 randomized controlled trial ('N-of-1 trial') of donepezil in the treatment of non-progressive amnestic syndrome. *Age Ageing* 2002;31:307-9.

29 Wegman AC, van der Windt DA, Bongers M, Twisk JW, Stalman WA, de Vries TP. Efficacy of temazepam in frequent users: a series of N-of-1 trials. *Fam Pract* 2005;22:152-9.

30 Träder J. Mutterkraut als prophylaktische phytotherapie bei migräne: N-of-1-studie zum individuellen wirksamkeitsnachweis [feverfew as preventive therapy for migraine: N-of-1-trial for verification of individual efficacy]. *ZFA* 2007;83:238-41.

31 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008;371:281-3.

32 Nonoyama ML, Brooks D, Guyatt GH, Goldstein RS. Effect of oxygen on health quality of life in COPD patients with transient exertional hypoxemia. *Am J Respir Crit Care Med* 2007;176:343-9.

33 Pereira JA, Holbrook AM, Dolovich L, Goldsmith C, Thabane L, Douketis JD, et al. Are brand-name and generic warfarin interchangeable? multiple n-of-1 randomized, crossover trials. *Ann Pharmacother* 2005;39:1188-93.

34 Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol* 2009;62:499-505.

35 Kamien M. The use of an N-of-1 randomised clinical trial in resolving therapeutic doubt. the case of a patient with an 'attention disorder'. *Aust Fam Physician* 1998;27(suppl 2):S103-5.

36 Sung L, Tomlinson GA, Greenberg ML, Koren G, Judd P, Ota S, et al. Serial controlled N-of-1 trials of topical vitamin E as prophylaxis for chemotherapy-induced oral mucositis in paediatric patients. *Eur J Cancer* 2007;43:1269-75.

37 Louly PG, Medeiros-Souza P, Santos-Neto L. N-of-1 double-blind, randomized controlled trial of tramadol to treat chronic cough. *Clin Ther* 2009;31:1007-13.

38 Wegman AC, van der Windt DA, de HM, Deville WL, Fo CT, de Vries TP. Switching from NSAIDs to paracetamol: a series of n of 1 trials for individual patients with osteoarthritis. *Ann Rheum Dis* 2003;62:1156-61.

39 Smith BJ, Appleton SL, Veale AJ, McElroy HJ, Veljkovic D, Saccoia L. Eformoterol n-of-1 trials in chronic obstructive pulmonary disease poorly reversible to salbutamol. *Chron Respir Dis* 2004;1:63-9.

40 Chin R, Chin RY, Lee BY. *Principles and practice of clinical trial medicine* . Academic Press, 2008.

41 Johannessen T, Petersen H, Kristensen P, Fosstvedt D, Kleveland PM, Dybdahl J, et al. Cimetidine on-demand in dyspepsia. experience with randomized controlled single-subject trials. *Scand J Gastroenterol* 1992;27:189-95.

42 Fedele DJ, Jones JA, Volicer L, Herz LR, Oppenheim FG. Effects of medication on parotid salivary flow rates in an individual with dementia of the alzheimer type. *Spec Care Dentist* 1993;13:215-8.

43 Coxeter PD, Schluter PJ, Eastwood HL, Nikles CJ, Glasziou PP. Valerian does not appear to reduce symptoms for patients with chronic insomnia in general practice using a series of randomised n-of-1 trials. *Complement Ther Med* 2003;11:215-22.

44 Nathan PC, Tomlinson G, Dupuis LL, Greenberg ML, Ota S, Bartels U, et al. A pilot study of ondansetron plus metopimazine vs. ondansetron monotherapy in children receiving highly emetogenic chemotherapy: a bayesian randomized serial N-of-1 trials design. *Support Care Cancer* 2006;14:268-76.

45 Irwig L, Glasziou P, March L. Ethics of n-of-1 trials. *Lancet* 1995;345:469.

46 Mahon JL, Feagan BG, Laupacis A. Ethics of n-of-1 trials. *Lancet* 1995;345:989.

47 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;63:1312-23

48 Punja S, Eslick I, Duan N, Vohra S, the DEcIDE Methods Center N-of-1 Guidance Panel. Chapter 2: An ethical framework for N-of-1 trials: clinical care, quality improvement, or human subjects research? In: Kravitz RL, Duan N, and the DEcIDE Methods Center N-of-1 Guidance Panel, eds. *Design and implementation of N-of-1 trials: a user's guide* . AHRQ Publication No 13(14)-EHC122-EF. Agency for Healthcare Research and Quality, 2014:13-22.

49 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;314:889-92.

50 Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687.

51 Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C. Reporting randomized, controlled trials of herbal interventions: An elaborated CONSORT statement. *Ann Intern Med* 2006;144:364-7.

52 Gagnier JJ, Boon H, Rochon P, Moher D, Barnes J, Bombardier C, et al. Recommendations for reporting randomized controlled trials of herbal interventions: explanation and elaboration. *J Clin Epidemiol* 2006;59:1134-49.

53 MacPherson H, Altman DG, Hammerschlag R, Li Y, Wu T, White A, et al. Revised STandards for reporting interventions in clinical trials of acupuncture (STRICTA): Extending the CONSORT statement. *Acupunct Med* 2010;28:83.

54 Boutron I, Moher D, Altman DG, Schulz K, Ravaud P, for the CONSORT group. Methods and processes of the CONSORT group: example of an extension for trials assessing nonpharmacologic treatments. *Ann Intern Med* 2008;148:W60-7.

55 Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295.

56 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;1:726-8.

57 Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PloS One* 2013;8:e66844.

58 Guyatt G, Montori V, Devereaux PJ, Schünemann H, Bhandari M. Patients at the center: in our practice, and in our use of language. *ACP J Club* 2004;140:A11-2.

59 Guyatt GH, Jaeschke R, Roberts R. N-of-1 randomized clinical trials in pharmacoepidemiology. In: Strom BL, ed. *Pharmacoepidemiology* . Wiley, 2005:665-80.

60 Gandhi GY, Murad MH, Fujiyoshi A, Mullan RJ, Flynn DN, Elamin MB, et al. Patient-important outcomes in registered diabetes trials. *JAMA* 2008;299:2543.

61 Paterson C. Measuring outcomes in primary care: a patient generated measure, MYMOP, compared with the SF-36 health survey. *BMJ* 1996;312:1016.

62 Paterson C, Allen JA, Browning M, Barlow G, Ewings P. A pilot study of therapeutic massage for people with parkinson's disease: the added value of user involvement. *Complement Ther Clin Pract* 2005;11:161-71.

63 Dziegielewski SF. Eye movement desensitization and reprocessing (EMDR) as a time-limited treatment intervention for body image disturbance and self-esteem: a single subject case study design. *J Psychother Independent Pract* 2000;1:1-16.

64 Johnston BC, Shamseer L, da Costa BR, Tsuyuki RT, Vohra S. Measurement issues in trials of pediatric acute diarrheal diseases: a systematic review. *Pediatrics* 2010;126:e222.

65 Sinha IP, Altman DG, Beresford MW, Boers M, Clarke M, Craig J, et al. Standard 5: selection, measurement, and reporting of outcomes in clinical trials in children. *Pediatrics* 2012;129(suppl 3):S146-52.

66 Miyar J, Adams CE. Content and quality of 10,000 controlled trials in schizophrenia over 60 years. *Schizophr Bull* 2013;39:226-9.

67 Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *Br J Psychiatry* 2000;176:249-52.

68 Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. *PLoS One* 2010;5:e9810.

69 Zwaigenbaum L, Dick P, Handley-Derry M, Malone M, Jacobson SE. "N of 1" trials of methylphenidate in two children with Williams syndrome and attention deficit hyperactivity disorder. *J Dev Phys Disabil* 2006;18:45-58.

70 Nikles J, Mitchell GK, Schluter P, Good P, Hardy J, Rowett D, et al. Aggregating single patient (n-of-1) trials in populations where recruitment and retention was difficult: the case of palliative care. *J Clin Epidemiol* 2011;64:471-80.

71 Cepeda MS, Acevedo JC, Alvarez H, Miranda N, Cortes C, Carr DB. An N-of-1 trial as an aid to decision-making prior to implanting a permanent spinal cord stimulator. *Pain Med* 2008;9:235-9.

72 Montori VM, Devereaux P, Adhikari NK, Burns KE, Eggert CH, Briel M, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA* 2005;294:2203.

73 Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, et al. Stopping randomized trials early for benefit and estimation of treatment effects. *JAMA* 2010;303:1180-7.

74 Guyatt GH, Briel M, Glasziou P, Bassler D, Montori VM. Problems of stopping trials early. *BMJ* 2012;344:e3863.

75 Nikles CJ, Yelland M, Glasziou PP, Del MC. Do individualized medication effectiveness tests (n-of-1 trials) change clinical decisions about which drugs to use for osteoarthritis and chronic pain? *Am J Ther* 2005;12:92-7.

76 Morrow KL, Fridriksson J. Comparing fixed- and randomized-interval spaced retrieval in anomia treatment. *J Commun Disord* 2006;39:2-11.

77 Schmid CH, Duan N, the DEcIDE Methods Center N-of-1 Guidance Panel. Chapter 4: Statistical design and analytic considerations for N-of-1 trials. In: Kravitz RL, Duan N, DEcIDE Methods Center N-of-1 Guidance Panel, eds. *Design and implementation of N-of-1 trials: a user's guide* . AHRQ Publication No 13(14)-EHC122-EF ed. Agency for Healthcare Research and Quality, 2014:33-53.

78 Woodfield R, Goodyear-Smith F, Arroll B. N-of-1 trials of quinine efficacy in skeletal muscle cramps of the leg. *Br J Gen Pract* 2005;55:181-5.

79 Price JD, Grimley EJ. N-of-1 randomized controlled trials ('N-of-1 trials'): singularly useful in geriatric medicine. *Age Ageing* 2002;31:227-32.

80 Arroll B, Miller P. Single case randomised trial of tryptophan vs placebo. *NZ Fam Physician* 1989;16:14-5.

81 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med* 1997;126:36-47.

82 Gage NA, Lewis TJ. Analysis of effect for single-case design research. *J Appl Sport Psychol* 2013;25:46-60.

83 Horner RH, Swaminathan H, Sugai G, Smolkowski K. Considerations for the systematic analysis and use of single-case research. *Educ Treat Child* 2012;35:269-90.

84 Parker RI, Vannest KJ, Davis JL. Effect size in single-case research: a review of nine nonoverlap techniques. *Behav Modif* 2011;35:303-22.

85 Heyvaert M, Wendt O, Van den Noortgate W, Onghena P. Randomization and data-analysis items in quality standards for single-case experimental studies. *J Special Educ* 2014; doi:0022466914525239.

86 Shadish WR. Statistical analyses of single-case designs the shape of things to come. *Curr Direct Psychol Sci* 2014;23:139-46.

87 Altman DG. Confidence intervals in practice. In: Altman DG, Machin D, Bryant TN, Gardner MJ, eds. *Statistics with confidence* . 2nd ed. BMJ Books, 2000:6-14.

88 Altman DG. Clinical trial and meta-analyses. In: Altman DG, Machin D, Bryant TN, Gardner MJ, eds. *Statistics with confidence* . 2nd ed. BMJ Books, 2000:120-38.

89 Bailar JC 3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. amplifications and explanations. *Ann Intern Med* 1988;108:266-73.

90 Zucker DR, Ruthazer R, Schmid CH, Feuer JM, Fischer PA, Kieval RI, et al. Lessons learned combining N-of-1 trials to assess fibromyalgia therapies. *J Rheumatol* 2006;33:2069-77.

91 Maggin DM, O'Keeffe BV, Johnson AH. A quantitative synthesis of methodology in the meta-analysis of single-subject research for students with disabilities: 1985–2009. *Exceptionality* 2011;19:109-35.

92 Gage NA, Lewis TJ. Hierarchical linear modeling meta-analysis of single-subject design research. *J Special Educ* 2012; doi:0022466912443894.

93 Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340:c221.

94 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264-9.

95 Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1-e34.

96 Haas DC, Sheehe PR. Dextroamphetamine pilot crossover trials and n of 1 trials in patients with chronic tension-type and migraine headache. *Headache* 2004;44:1029-37.

97 Senn S. *Cross-over trials in clinical research* . 2nd ed. Wiley, 2002.

98 Maggin DM, Swaminathan H, Rogers HJ, O'Keeffe BV, Sugai G, Horner RH. A generalized least squares regression approach for computing effect sizes in single-case research: application examples. *J School Psychol* 2011;49:301-21.

99 Rochon J. A statistical model for the "N-of-1" study. *J Clin Epidemiol* 1990;43:499-508.

100 Parker RI, Hagan-Burke S, Vannest K. Percentage of all non-overlapping data (PAND) an alternative to PND. *J Special Educ* 2007;40:194-204.

101 Johnston MV, Smith RO. Single subject designs: current methodologies and future directions. *OTJR* 2010;30:4.

102 Gordon K, Bawden H, Camfield P, Mann S, Orlik P. Valproic acid treatment of learning disorder and severely epileptiform EEG without clinical seizures. *J Child Neurol* 1996;11:41-3.

103 Dales RE, Lunau MA, Tierney MG, Schweitzer I. Chronic cough responsive to ibuprofen. *Pharmacotherapy* 1992;12:331-3.

104 Estrada CA, Young MJ. Patient preferences for novel therapy: an N-of-1 trial of garlic in the treatment for hypertension. *J Gen Intern Med* 1993;8:619-21.

105 Lehman R, Loder E. Missing clinical trial data. *BMJ* 2012;344:d8158.

106 AllTrials. All trials registered: all results reported. www.alltrials.net.

107 Avins AL, Bent S, Neuhaus JM. Use of an embedded N-of-1 trial to improve adherence and increase information from a clinical study. *Contemp Clin Trials* 2005;26:397-401.

108 Reed JF III. AB/BA crossover trials-binary outcome. *J Mod App Stat Meth* 2005;5:19.

109 Johannessen T, Kristensen P, Petersen H, Fosstvedt D, Loge I, Kleveland PM, et al. The symptomatic effect of 1-day treatment periods with cimetidine in dyspepsia. combined results from randomized, controlled, single-subject trials. *Scand J Gastroenterol* 1991;26:974-80.

110 Yelland MJ, Nikles CJ, McNairn N, Del Mar CB, Schluter PJ, Brown RM. Celecoxib compared with sustained-release paracetamol for osteoarthritis: a series of n-of-1 trials. *Rheumatology (Oxford)* 2007;46:135-40.

111 Van RR, Bayley M, Garner S, Burke IM, Fawcett S, Hart A, et al. N of 1 study: amantadine for the amotivational syndrome in a patient with traumatic brain injury. *Brain Inj* 1995;9:49-53.

112 Hoffmeyer CA. A multiple single-case study experimental design exploring the effect of therapeutic touch on women with migraine headaches [dissertation]. University of Colorado Health Sciences Center, 2000.

113 Nikles CJ, Mitchell GK, Del Mar CB, Clavarino A, McNairn N. An n-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention-deficit/hyperactivity disorder. *Pediatrics* 2006;117:2040-6.

114 Tison F, Nègre-Pagès L, Meissner WG, Dupouy S, Li Q, Thiolat M, et al. Simvastatin decreases levodopa-induced dyskinesia in monkeys, but not in a randomized, placebo-controlled, multiple cross-over ("n-of-1") exploratory trial of simvastatin against levodopa-induced dyskinesia in parkinson's disease patients. *Parkinsonism Relat Disord* 2013;19:416-21.

115 De Angelis C, Drazen JM, Frizelle PFA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the international committee of medical journal editors. *N Engl J Med* 2004;351:1250-1.

116 Food and Drug Administration Amendment Act of 2007, Title VIII, Section 801. 2007). See: www.fda.gov/regulatoryinformation/legislation/federalfooddrugandcosmeticactfdcact/significantamendmentstothefdcact/foodanddrugadministrationamendmentsactof2007/default.htm.

117 Paz-Filho GJ, Babikian T, Asarnow R, Delibasi T, Esposito K, Erol HK, et al. Leptin replacement improves cognitive development. *PLoS One.* 2008;3:e3098.

118 Chan AW, Hróbjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.

119 Moher D. Reporting research results: a moral obligation for all researchers. *Can J Anaesth* 2007;54:331-5.

120 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Obstet Gynecol* 2009;114:1341-5.

121 Carry over. Glossary of Cochrane terms 2013. www.cochrane.org/glossary/5#term137.

122 Campbell MJ, Machin D, Walters SJ. *Medical statistics: a textbook for the health sciences* . John Wiley, 2010.

Tables

Table 1| CENT 2015 checklist*; CONSORT 2010 checklist items with modifications or additions for individual or series of N-of-1 trials; empty items in the CENT 2015 column indicate no modification from the CONSORT 2010 item

Section/Topic	CONSORT 2010 No	Item	CENT 2015 No	Item
Title and abstract				
	1a	Identification as a randomised trial in the title	1a	Identify as an "N-of-1 trial" in the title
				For series: Identify as "a series of N-of-1 trials" in the title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1b	For specific guidance, see CENT guidance for abstracts (table 2)
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	2a.1	
			2a.2	Rationale for using N-of-1 approach
	2b	Specific objectives or hypotheses	2b	
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	3a	Describe trial design, planned number of periods, and duration of each period (including run-in and wash out, if applicable)
				In addition for series: Whether and how the design was individualized to each participant, and explain the series design
	3b	Important changes to methods after trial start (such as eligibility criteria), with reasons	3b	
Participant(s)	4a	Eligibility criteria for participants	4a†	Diagnosis or disorder, diagnostic criteria, comorbid conditions, and concurrent therapies.
				For series: Same as CONSORT item 4a
	4b	Settings and locations where the data were collected	4b†	
			4c	Whether the trial(s) represents a research study and if so, whether institutional ethics approval was obtained
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	5	The interventions for each period with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	6a.1	
			6a.2	Description and measurement properties (validity and reliability) of outcome assessment tools
	6b	Any changes to trial outcomes after the trial commenced, with reasons	6b	
Sample size	7a	How sample size was determined	7a	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	7b	
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence	8a	Whether the order of treatment periods was randomised, with rationale, and method used to generate allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	8b	When applicable, type of randomisation; details of any restrictions (such as pairs, blocking)
			8c	Full, intended sequence of periods
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	9	

Table 1 (continued)

Section/Topic	CONSORT 2010 No	CONSORT 2010 Item	CENT 2015 No	CENT 2015 Item
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	10	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	11a	
	11b	If relevant, description of the similarity of interventions	11b	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12a	Methods used to summarize data and compare interventions for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12b	*For series:* If done, methods of quantitative synthesis of individual trial data, including subgroup analyses, adjusted analyses, and how heterogeneity between participants was assessed *(for specific guidance on reporting syntheses of multiple trials, please consult the PRISMA Statement)*
			12c	Statistical methods used to account for carryover effect, period effects, and intra-subject correlation
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	13a.1	Number and sequence of periods completed, and any changes from original plan with reasons
			13a.2	*For series:* The number of participants who were enrolled, assigned to interventions, and analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons	13c	*For series:* Losses or exclusions of participants after treatment assignment, with reasons, and period in which this occurred, if applicable
Recruitment	14a	Dates defining the periods of recruitment and follow-up	14a†	
	14b	Why the trial ended or was stopped	14b	Whether any periods were stopped early and/or whether trial was stopped early, with reason(s).
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	15†	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	16	For each intervention, number of periods analysed. *In addition for series:* If quantitative synthesis was performed, number of trials for which data were synthesized
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	17a.1	For each primary and secondary outcome, results for each period; an accompanying figure displaying the trial data is recommended.
			17a.2	For each primary and secondary outcome, the estimated effect size and its precision (such as 95% confidence interval) *In addition for series:* If quantitative synthesis was performed, group estimates of effect and precision for each primary and secondary outcome
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	17b	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	18	Results of any other analyses performed, including assessment of carryover effects, period effects, intra-subject correlation *In addition for series:* If done, results of subgroup or sensitivity analyses
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	19	All harms or unintended effects for each intervention. *(for specific guidance see CONSORT for harms)*
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	20	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	21	

Table 1 (continued)

| Section/Topic | CONSORT 2010 | | | CENT 2015 | |
	No	Item		No	Item
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		22	
Other information					
Registration	23	Registration number and name of trial registry		23	
Protocol	24	Where the full trial protocol can be accessed, if available		24	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		25	

*It is strongly recommended that this checklist be read in conjunction with the CENT 2015 Explanation and Elaboration[24] for important clarification on the items. The copyright for CENT (including checklist) is held by the CENT Group and is distributed under a Creative Commons Attribution (CC-BY 4.0) license.

†Caution should be taken when reporting potentially identifying information pertaining to CENT items 4a, 4b, 14a, and 15.

Table 2| CENT abstract considerations (modifications or additions to CONSORT Statement for Abstracts)

| Item	CONSORT for abstracts	Extension for N-of-1 designs
Title	Identification of the study as randomised	Identification of the study as an N-of-1 trial or series of N-of-1 trials in the title
Authors*	Contact details for the corresponding author	
Trial design	Description of trial design (such as parallel, cluster, non-inferiority)	Description of trial design, number of periods, and period duration
Methods:		
Participant(s)	Eligibility criteria for participants and the settings where the data were collected	For individual trial, clinical condition under study For series, eligibility criteria for participants
Interventions	Interventions intended for each group	Interventions intended for each period
Objective	Specific objective or hypothesis	
Outcome	Clearly defined primary outcome for this report	
Randomisation	How participants were allocated to interventions	
Blinding (masking)	Whether participant(s), care givers, and those assessing the outcomes were blinded to group assignment	
Results:		
Numbers randomised	Number of participants randomised to each group	For individual N-of-1 trial, the number and sequence of periods completed For series, number of individual trials carried out
Recruitment	Trial status	Not applicable
Numbers analysed	Number of participants analysed in each group	For individual N-of-1 report, number of periods analysed for each intervention For series, the number of participants analysed
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	
Harms	Important adverse events or side-effects.	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register, if applicable	
Funding	Source of funding	

*For conference abstracts.

Figures
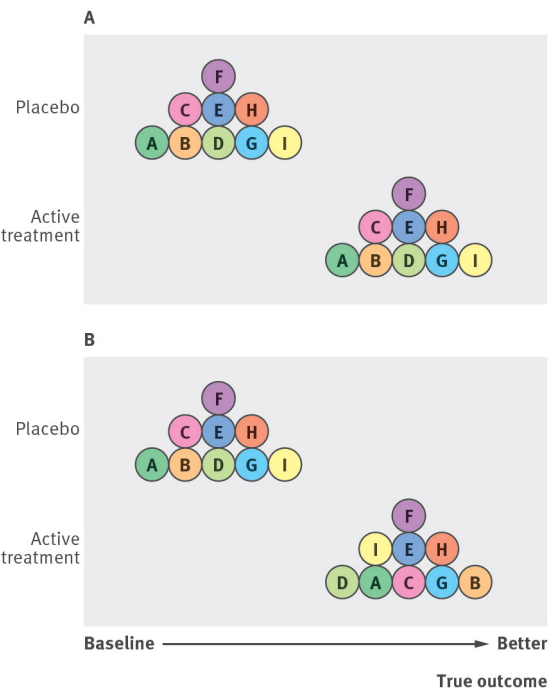
Fig 1 Illustration of how overall benefit at the group level may not benefit individual patients equally. Panel A shows no interaction between patients and treatment, where all individuals improve by the same amount; panel B shows interaction between patients and treatment such that each patient improves by an individual amount
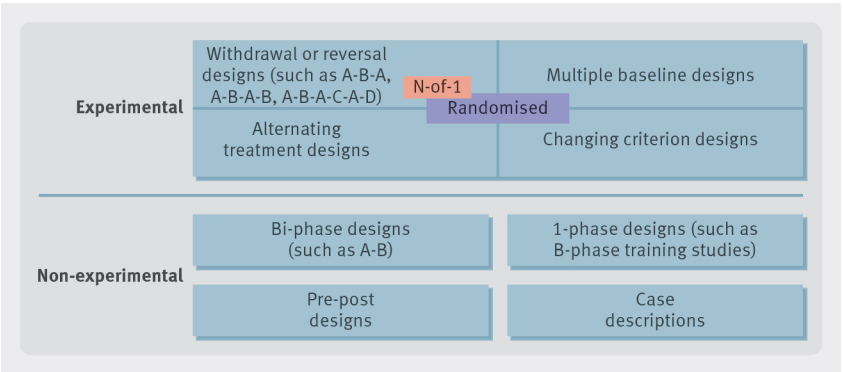

Fig 2 Common single case designs. CENT is applicable to a subset of the "Withdrawal/reversal designs" category, which may or may not include the use of randomisation, designated by the red "N-of-1" box (adapted from [17])
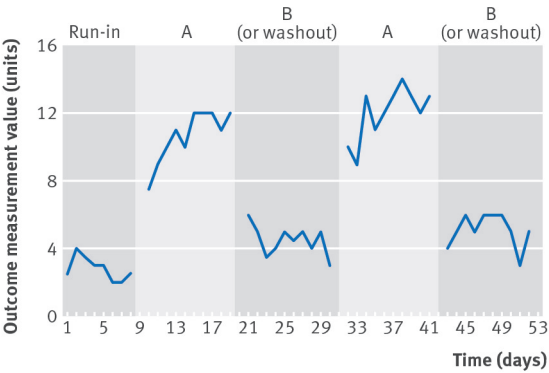
Fig 3 N-of-1 trial pictorial; suggested visual representation of data from an individual N-of-1 trial
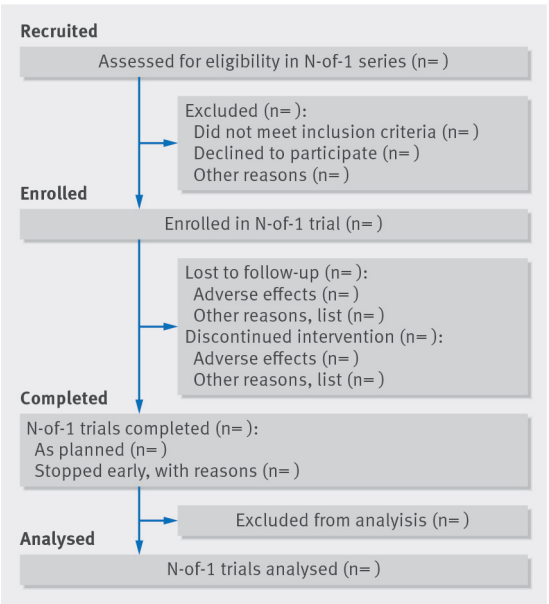
Fig 4 CENT flow diagram; suggested representation of the flow of participants in a series of N-of-1 trials
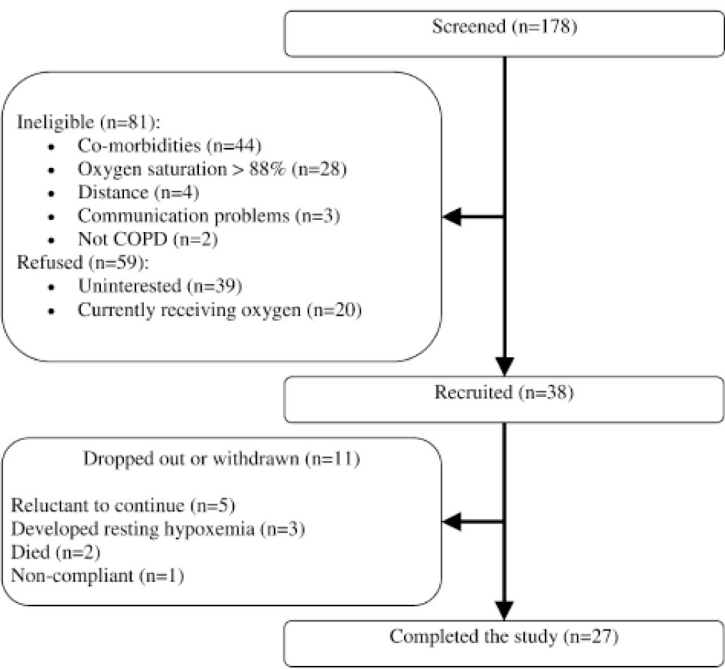
**Fig 5** Flow diagram from Nonoyama et al[32]

Table 1 Baseline characteristics (mean of Visit 1 and 2)

Subjects (male/female)	27 (21/6)
Age SD (years)	70.4 (7.9)
Current Smoker n (%)	5 (18.5)
Pre salbutamol lung function	Mean (SD)
FEV$_1$ litres	1.03 (0.34)
FEV$_1$ % predicted	36.1 (10.2)
FVC litres	2.77 (0.87)
Post salbutamol lung function	
FEV$_1$ litres	1.15 (0.39)
FEV$_1$ reversibility % baseline	12.1 (6.9)
Reversibility % predicted	3.9 (2.5)
FVC litres	3.06 (0.88)
FEV$_1$/FVC	38.0 (7.1)
Six minute walk distance metres[1]	316 (117)
HRQL-CRQ	Mean (SD)
Dyspnoea	3.3 (1.09)
Emotion	4.7 (0.93)
Fatigue	3.4 (1.14)
Mastery	4.6 (1.0)
Symptom scores[1]	
Day time	2.4 (0.86)
Night time	1.1 (0.82)
Concurrent medications	Mean (SD)
SABA	27
Daily doses[2]	3.82 (1.8)
Ipratropium bromide	18
Daily doses[2]	3.84 (1.3)
Inhaled corticosteroids	26
Oral corticosteroids	7
LABA[3]	5
Theophylline	5

[1]Baseline six minute walk distance and symptom scores based on 25 and 24 subjects respectively.
[2]Based upon run in period medication use: diary records completed for 24 (SABA) and 17 (ipratropium bromide). 1 dose SABA = 2 puffs; 1 dose ipratropium bromide = 4 puffs.
[3]Withdrawn prior to Visit 2.

Fig 6 Table of baseline characteristics from Smith et al[39]
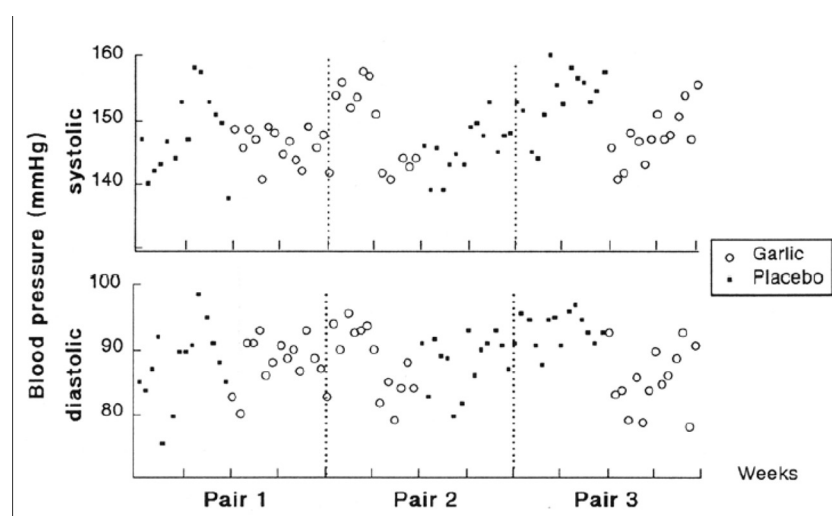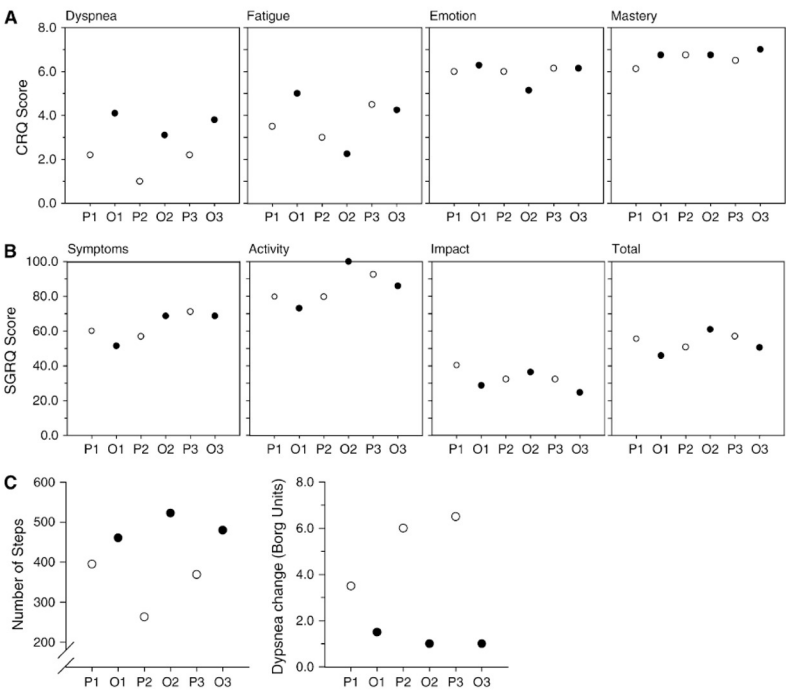

Fig 7 Results from Estrada et al[104]

Fig 8 Results from Nonoyama et al[32]

Table 1

Blood pressure response data in N-of-1 substudy

Measurement time	Blood pressure outcomes		
	Difference (mm Hg)*	95% confidence interval	*p*-value
Morning systolic BP	7.51	−1.20 to 16.22	0.09
Morning diastolic BP	2.06	−1.21 to 5.33	0.21
Evening systolic BP	−3.78	−7.54 to −0.02	0.05
Evening diastolic BP	3.45	−0.87 to 7.77	0.11

* Difference between adjusted mean blood pressures in study phases (positive values denote higher mean blood pressures in study-medicine periods compared to known placebo periods).

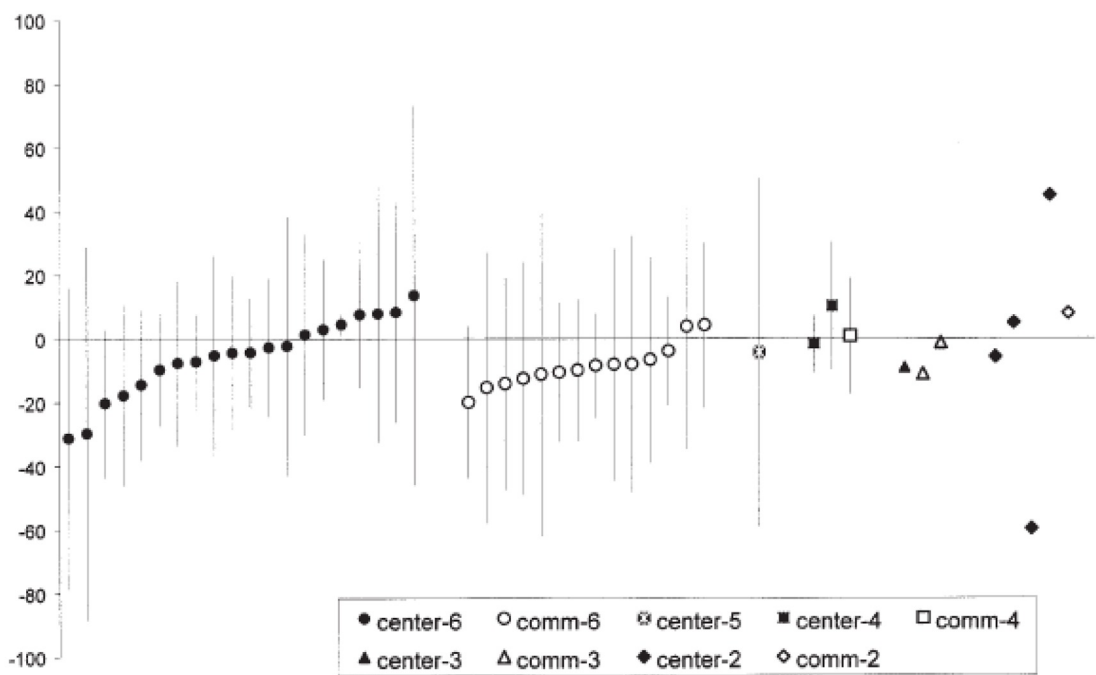**Fig 9** Individual trial data from Avins et al[107]

Figure 2. Individuals' treatment response variations shown as differences in mean FIQ scores taking AMT+FL and AMT with 95% confidence intervals. Center: center-based practice; comm: community-based practice.

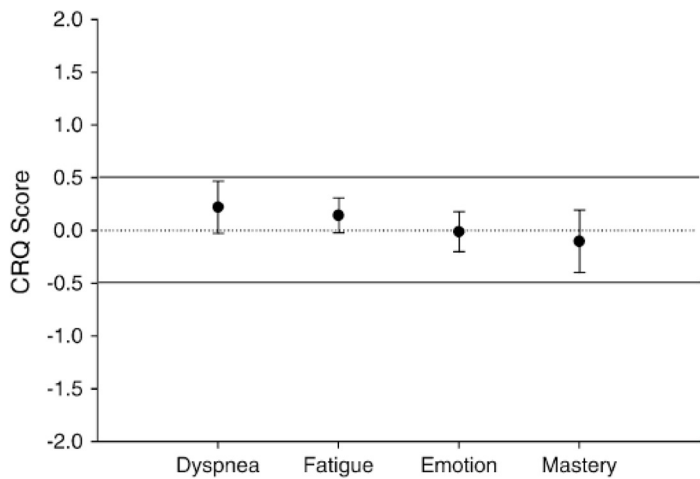**Fig 10** Individual trial data in a series from Zucker et al[90]

Fig 11 Series data from Nonoyama et al[32]

Table 3   Overall group estimates of the proportion of treatment success, and associated 95% confidence interval (95% CI), to valerian for the six sleep outcome variables

Sleep variables	Proportion of treatment success	95% CI
Latency to sleep onset	0.43	(0.29, 0.57)
Number of night awakenings	0.35	(0.23, 0.47)
Total sleep time	0.35	(0.23, 0.47)
Sleep quality	0.49	(0.35, 0.63)
Post-slumber refreshment in the morning	0.40	(0.28, 0.52)
Energy level in previous day	0.55	(0.43, 0.67)

Fig 12 Series data from Coxeter et al[43]