Title

Reporting quality of N-of-1 trials published between 1985 and 2013: a systematic review

Authors

Li, Jiang
Gao, Wei
Punja, Salima
et al.

Publication Date

2016-08-01

DOI

10.1016/j.jclinepi.2015.11.016

Peer reviewed

Reporting Quality of N-of-1 Trials Published between 1985 and 2013: A Systematic Review

Jiang Li[1,2,3], Wei Gao[4], Salima Punja[5], Bin Ma[1,2], Sunita Vohra[6], Naihua Duan[7], Nicole Gabler[8], Kehu Yang[1,2*], Richard L. Kravitz[9*]

1. Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China.
2. Key Laboratory of Evidence Based Medicine and Knowledge Translation of Gansu Province, Lanzhou, China.

3. Cancer Hospital, Chinese Academy of Medical Sciences, Beijing, China;

4. Institute of Biophysics, Chinese Academy of Sciences, Beijing, China;

5. Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada;

6. Complementary and Alternative Research and Education (CARE) Program, Department of Pediatrics and School of Public Health, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada;

7. Departments of Psychiatry and Biostatistics, Columbia University, New York, United States;

8. Center for Clinical Epidemiology and Biostatistics and Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, United States;

9. Department of Internal Medicine and Center for Healthcare Policy and Research, University of California, Davis, United States.

Corresponding Authors

Richard Kravitz: rlkravitz@ucdavis.edu

Kehu Yang: yangkh2006@163.com

ABSTRACT

Objective: To evaluate the quality of reporting of single-patient (N-of-1) trials published in the medical literature based on the CONSORT Extension for N-of-1 Trials (CENT) statement and to examine factors that influence reporting quality in these trials.

Methods: Through a search of 10 electronic databases, we identified N-of-1 trials in clinical medicine published between 1 January 1985 and 31 December 2013. Two reviewers screened articles for eligibility and independently extracted data. Quality assessment was performed using the CENT statement. Discrepancies were resolved by consensus.

Results: We identified 112 eligible N-of-1 trials published in 87 journals and involving a total of 2278 patients. Overall agreement between the two evaluators for compliance with CENT criteria was 0.80 (95% CI, 0.79 to 0.82). Trials assessed pharmacology and therapeutics (87%), behavior (11%), or diagnosis (2%). Although 87% of articles described the trial design (including the planned number of subjects and length of treatment period), the median percentage of specific CENT elements in the Methods was 41% (range 16% to 87%) and the median percentage in the Results was 38% (range 32% to 93%). First authors were predominantly from North America (46%), Europe (29%), and Australia (17%). Quality of reporting was higher in papers published in journals with relatively high impact factors (P=0.004).

Conclusion: The quality of reporting of published N-of-1 trials is variable and needs improvement. Because the CENT guidelines were not published until near the end of the period of this review, these results represent a baseline from which improvement may be expected in the future.

Key words: Single-patient (N-of-1) trials; CONSORT guidelines; reporting quality

1. INTRODUCTION

Individual-patient, or N-of-1, trials are prospective, multiple-period, cross-over studies conducted with a single patient [1, 2]. N-of-1 trials are often published as a series in which the same intervention-condition pair is assessed in multiple patients [3]. With the increasing recognition that, for certain conditions (chronic, symptomatic, non-fatal) and certain treatments (relatively short duration of action), N-of-1 trials can provide high-quality evidence for clinical care has come the need to report these trials to the highest standards. [4] Although trial quality cannot be assessed directly [5], reporting quality is considered a correlate of and arguably, an imperfect but serviceable proxy for—study trial quality [6]. If aspects of study conduct are not reported, it can be difficult to assess the quality of the study [7-10]. Reporting quality —including the rationale of the study, the trial design and activities, measurements, and analysis—is the basis for both peer reviewers' and readers' judgments of scientific merit and clinical applicability [11].

The CONSORT Statement provides researchers with guidance on clear and accurate reporting of randomized trials [7]. The recently published CONSORT Extension for N-of-1 Trials (CENT) provides important guidance on reporting N-of-1 trials, accounting for the unique features of these single-patient crossover studies [12]. Furthermore, the CENT statement serves as a checklist to help other researchers assess the validity and credibility of the trial and readers understand the trial better.

We assessed the quality of reports of N-of-1 trials published between 1985 and 2013 by determining their degree of adherence to the CENT criteria. The goal of the review was to describe the current state of reporting quality of N-of-1 trials and to investigate factors associated with better reporting quality.

2. Methods

2.1 Search strategy

The search strategy sought to identify N-of-1 trials published between January 1, 1985, and December 31, 2013. The 1985 start date was chosen to account for the first major wave of N-of-1 trials conducted in clinical medicine [13,14]. Records were identified from the Cochrane Central Register of Controlled Trials, PubMed, EMBASE, and ISI Web of Knowledge, as well as clinicaltrials.gov, the Chinese Clinical Trial Registry (ChiCTR) (http://www.chictr.org), the Chinese Biomedicine Literature Database (CBM), Wanfang Database, Chinese Scientific Journal Full-text Database (CSJD), Oriental Medicine Advanced Searching Integrated System (OASIS) (a Korean Database), and Japanese Institutional Repositories Online (JAIRO). Where possible, we also examined published, indexed conference proceedings; books; ongoing studies from ClinicalTrials.gov or ChiCTR; and theses and dissertations. Ancestry searches (examining reference lists in selected articles to find other relevant studies) were also carried out when possible. Published N-of-1 protocols were also screened but were ultimately excluded.

The search strategies were developed by one of the co-authors (Ma B) who has

worked on medical literature searches for 7 years.

2.2 Eligibility criteria

Included articles 1) were published in the medical literature between January 1985 through December 2013 in English or Mandarin Chinese (the two languages known to the investigators); 2) pertained to either clinical or behavioral topics; 3) reported single or combined N-of-1 trials involving at least two interventions over at least two pairs (i.e. ABAB); and 4) described allocation to treatment periods with terms such as “random,” “randomly,” “randomized,” or randomization.”

Studies were excluded if they were: 1) case reports; 2) protocols of N-of-1 trials; 3) systematic reviews of individual patient data; or 4) reviews and editorials on N-of-1 trials. Publications reporting results from the same participants were excluded.

2.3 Screening

Two reviewers independently screened the titles and abstracts of identified studies for initial eligibility. Two reviewers independently screened the full-text articles according to the inclusion and exclusion criteria described above. Disagreements were resolved in consultation with the third party.

2.4 Data Extraction

Two reviewers independently extracted data from each article using two data extraction forms. The first extraction form captured *publication* characteristics: first author’s name, year of publication, geographical region where the trial took place or

the region in which the first author resided, medical journal indexing information by Thomson Reuter's Journal Citation Reports (http: //thomsonreuters.com/products_services/science/free/essays/impact_factor/), the impact factor of the journal and medical content area.

The second form was for assessing quality of reporting using CENT guidelines (Table 2). The two reviewers were trained conjointly to apply the Statement by an experienced systematic reviewer.

As part of their training, reviewers received the CONSORT Statement for Reporting Randomized Trials, which provides the definitions and rationale for each checklist item and examples of good reporting [15]. All reviewers then received a confidential pre-publication version of the CENT Statement with instructions on how to evaluate N-of-1 trials. Both reviewers (BM and JL) are physician-investigators who have worked on reporting guidelines for at least 4 years.

2.5 Rating of Overall Reporting Quality

Publication characteristics are described with counts and percentages. The impact factor of the publishing journal was determined from the Science Citations Index [16]. High-impact journals were defined as those with impact factors above the median of the journals included in the study, and low-impact journals had impact factors below the median. Journals not indexed by Thomson-Reuters were assigned an impact factor of 0. Articles were also classified geographically by country of residence of the first author.

Overall reporting quality was assessed using the 43-item CENT Statement for series of N-of-1 trials and the 40-item CENT Statement for single N-of-1 trials (items 12b, 14b, and 14c are for series of N-of-1 trials). Each item was marked as “yes” if it was fully reported (Y, scored as 1 point) and “no” if it was not clearly reported or not definitely stated (N, scored as 0 points). Scores on both scales were then transformed to percentages based on the relevant denominator (i.e., 43 or 40). Percentage scores were categorized as: excellent (>90%), good (between 50% and 90%), and poor (<50%). Each reviewer was blinded to the other's ratings. Inter-rater agreement was assessed with the kappa statistic.

2.6 Statistical Analysis

All information was managed using Microsoft Excel 2010 Software. Inter-rater agreements were obtained independently for each item with the Cohen’s kappa (κ) statistic[17]. Continuous variables are summarized as frequencies, means, and standard deviations.

Univariate analysis was used to compare the mean overall CENT scores for each subgroup (such as high- and low-impact factor group and year of publication by decade). Alpha was set at 0.05, and all tests were two-tailed. All statistical analyses were performed with the SAS software, version 9.2.

Results

3.1 Search Results and Trial Characteristics

Of 756 potentially relevant articles, 112 (including 6 in Chinese) met the inclusion

criteria (Figure 1). Of the 112 trials, a plurality were published as a series (65%), were conducted in North America (47%), and assessed pharmaceutical interventions (87%) (Table 1). The median impact factor was 2.36 (range, 0.36 to 14).

3.2 Ratings of Overall Quality and Factors Associated with Reporting Quality

Inter-rater agreement between the two evaluators averaged 0.80 (95% CI, 0.79 to 0.82) but ranged from 0.63 (specific elements of N-of-1 trial design, including planned sample size and duration of each period) to 1.0 (table describing baseline characteristics of participants). The mean (SD) CENT checklist score was 16.8 (5.3) out of 40 (42%) for single-patient trials and 25.7 (9.4) out of 43 (60%) for combined trials. The quality of reporting was excellent (score >90% of CENT checklist items) in 12 (10%) trials, good (50%-90%) in 56 (50%), and poor (<50%) in 44 (39%).

There was considerable item-by-item variation in the likelihood that CENT criteria would be met (Table 2). For example, the nature of the interventions compared and the flow of participants were adequately described for at least 80% of studies and sample size calculations were reported in 87%, whereas randomization method were described in only 29% . Overall, a majority items in the *Methods* and in the *Results* were included in less than half the articles.

Mean reporting scores for the 70 trials published after 2000 were significantly higher than those for earlier trials ($P<0.001$). Mean scores were significantly higher for trials published in higher impact journals compared with those in journals of lower or unknown impact ($P=0.004$). Scores did not significantly differ by the geographic

origin of the first author's home institution (see Table 3)

Discussion

The results of this comprehensive review support three principal conclusions. First, the rate of publication of N-of-1 trials, while still relatively infrequent, appears to be increasing over time. Second, CENT criteria for reporting quality of N-of-1 trials can, for the most part, be assessed reliably, although inter-rater agreement was greater for some elements than others like random assignment, blind method and results. Third, N-of-1 trial reporting quality is highly variable, leaving substantial room for improvement.

After emerging in the clinical literature to much fanfare [1, 2, 18], N-of-1 trials appeared in the literature sporadically during the late 1990s and then more consistently through the first decade of the current century. The apparent acceleration starting in 2010 may reflect a nexus between increased interest in personalized medicine and methods to enhance the delivery of patient-centered care [19].

The CENT guidelines for N-of-1 trial reporting were developed through a rigorous modified Delphi process involving expert input [12]. Our results indicate that inter-rater agreement for most CENT criteria was good-to-excellent, but a few important elements had suboptimal kappa scores. For example, item 3a (kappa, 0.63) asks for key elements of trial design. This information is critical for readers. Future iterations of CENT may want to split out the various sub-domains of item 3a into separate items, each accompanied by clear definitions. On the other hand, the similar

result for item 7a (how sample size was determined) may simply reflect insufficient reviewer training.

In assessing the landscape of N-of-1 trial reporting over 27 years, our results support two disparate, but not necessarily contradictory, conclusions. On the one hand, the majority of CENT reporting items was met by at least 50% of articles reviewed. Of particular importance for assessing the likelihood of biased conclusions, most articles reported on trial design (87%), sequence generation (79%), interventions (87%), and pre-specified outcomes (61%). However, only 23% reported on allocation concealment, 29% on method of randomization, and 33% on the intended sequence of periods. Each of these deficits could bias conclusions and influence interpretation [20]. There were also substantial deficits in reporting of baseline characteristics of participants (36%), outcome differences by treatment period (32%), ethics approval (39%), and funding source (36%). However, authors could not have been expected to provide information on funding source until 2001, when the International Committee of Medical Journal Editors published a committee-written editorial urging that authors be required to divulge potential financial conflicts of interest [21].

Our analysis also suggests that (ignoring the dip since 2010) the quality of N-of-1 trial reporting may be improving over time. We might expect reporting to improve further after publication of the CENT guidelines.

Finally, we found reporting quality to be marginally better in journals with a relatively high impact factor. This is not surprising because higher impact journals

may have broader reviewer networks to review and edit manuscripts, more statistical consulting resources, and more rigorous methodological standards.

Limitations of the Study

A limitation of the study was that publication of the CENT statement followed the publication of the reports it was used to assess. However, as journals begin to adopt the CENT statement, experience with the CONSORT guidelines provides grounds for optimism: in a systematic review by Turner et al., journal endorsement of the CONSORT guidelines was associated with more complete scientific reporting [22].

We included only articles in English and Chinese, because of language limitations. As such, we may not have captured otherwise eligible N-of-1 trials published in other languages. The databases chosen were inclusive of most but not all relevant journals. Thus, some trials were probably missed, but the total number of missed trials was likely to be small. We assessed each item with a ‘‘yes’’ or ‘‘no’’ response according to whether the author had reported the detail contents listed in the refined items. A quantitative score to assess each article might have provided greater discrimination [6, 22]. The CENT scoring system used in this article gave equal (and therefore arbitrary) weights to each item. Assigning weights on a different basis could be difficult in the absence of consensus on the relative importance of the different items [23]. Finally, although many readers of scientific literature assume that publication in a high impact journal is a proxy for greater quality [24] the SCI

database covers thousands of journals whose impact factors may vary considerably from year-to-year and can be subject to gaming.

Conclusions

N-of-1 trials are a potentially powerful method for more informed, individualized decision making. However, the utility and influence of n-of-1 trials will depend heavily on the rigor of their design, conduct, and reporting. As judged by CENT criteria, N-of-1 trials published between 1985 and 2013 display a high degree of variability in their reporting quality. As a developing patient- oriented trial methodology, N-of-1 trials must be adequately reported, and CENT should help improve the quality of these reports. Better reporting would serve the interests of researchers, reviewers, journal editors, clinicians, and patients.

Acknowledgement

The authors would like to thank Larissa Shamseer for sharing drafts of the CENT criteria and supporting documents.

REFERENCES

1. Guyatt GH, 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.

2. Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Ann Intern Med 1990; 112: 293–9.

3. Punja S, Bukutu C, Shamseer L, Sampson M, Hartling L, Urichuk L, Vohra S. The design, analysis and meta-analysis of N-of-1 trials: A tapestry of heterogeneity. Submitted to Journal of Clinical Epidemiology. 2015

4. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998; 352: 609-13.

5. Mhaskar R, Djulbegovic B, Magazin A, Soares HP, Kumar A. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. J Clin Epidemiol 2012; 65: 602-9.

6. Mayo-Wilson E, Grant S, Hopewell S, Macdonald G, Moher D, Montgomery P. Developing a reporting guideline for social and psychological intervention

trials. Trials. 2013; 14: 242.

7. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 1996; 276: 637-9.

8. Moher D, Altman DG, Schulz KF, Elbourne DR. Opportunities and challenges for improving the quality of reporting clinical research: CONSORT and beyond. CMAJ 2004; 171: 349-50.

9. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. JAMA 1994; 272: 125-8.

10. Altman DG, Moher D, Schulz KF. Improving the reporting of randomised trials: the CONSORT Statement and beyond. Stat Med 2012; 31:2985-97.

11. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM Statement. Onkologie 2000; 23: 597-602.

12. Shamseer L, Sampson M, Bukutu C, Barrowman N, Altman D, Moher D, et al. CONSORT extension for N-of-1 trials (CENT) guidelines. BMC Complement Altern Med 2012; 12: 410.

13. Keller JL, Guyatt GH, Roberts RS, Adachi JD, Rosenbloom D. An N of 1 service:

applying the scientific method in clinical practice. Scand J Gastroenterol Suppl 1988; 147: 22-9.

14. Guyatt G, Sackett D, Adachi J, Roberts R, Chong J, Rosenbloom D, Keller J. A clinician's guide for conducting randomized trials in individual patients. CMAJ 1988; 139: 497-503. 15. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010; 63: 834-40.

16. http://wokinfo.com/essays/impact-factor/

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-74.

18. Larson EB, Ellsworth AJ, Oas J. Randomized clinical trials in single patients during a 2-year period. JAMA 1993; 270(22): 2708-12.

19. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1clinical trial: the ultimate strategy for individualizing medicine? Per Med 2011; 8(2): 161-73.

20. Dwan K, Gamble C, Williamson PR, Kirkham JJ; Reporting Bias Group. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One 2013; 8: e66844.

21. Blum JA, Freeman K, Dart RC, Cooper RJ. Requirements and definitions in conflict of interest policies of medical journals. JAMA. 2009; 302(20):

2230-4.

22. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, Dias S, Schulz KF, Plint AC, Moher D. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev 2012; 11: MR000030.

23. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? BMJ 2008; 336: 1472–4.

24. Björk BC, Solomon D. Open access versus subscription journals: a comparison of scientific impact. BMC Med 2012; 10: 73.

Table 1. Characteristics of Included Articles

Characteristic	n (%)
Year of publication	
1985-1989	1 (1)
1990-1999	42 (38)
2000-2009	45 (40)
2010-2012	24 (21)
First author's region of residence	
North America	52 (47)
Europe	33 (29)
Australia	19 (17)
Asia	7 (6)
South America	1 (1)
Impact factor of journal	
Missing (impact factor=0)	18 (16)
Below the median value of 2.36	40 (36)
Above the median value of 2.36	54 (48)
Number of including patients	
Single	39 (35)
Series	73 (65)
Trial topic	
Pharmacology and therapeutics	98 (87)
Behavior	12 (11)
Diagnosis	2 (2)
Total articles reviewed	112 (100)

Table 2. Quality of Reporting of 112 N-of-1 Trials as Assessed by Adherence to the CONSORT Extension for N-of-1 Trials (CENT) Statement

Section/Topic	Item number and description	Item present, %	95% CI	Agreement between reviewers, Kappa
	Title and abstract			
	1a. Identify as N-of-1 trial in the title	72	0.67-0.76	0.97
	1b. Structured summary of trial design, methods, results, and conclusions*	56	0.51-0.61	0.98
	Introduction			
Background and objectives	2a. Scientific background and explanation of rationale	86	0.82-0.89	0.87
	2b. Specific objectives or hypotheses	84	0.80-0.87	0.71
	2c. Provide rationale for using N-of-1 approach	57	0.52-0.62	0.84
	Methods			
Trial design	3a. Describe the trial design, planned number, and duration of each period (including run-in and wash out, if applicable) with rationale. In addition for series: Whether and how the design was individualized to each participant, and explanation of the series design	87	0.84-0.91	0.63
	3b. Important changes to methods after trial commencement (such as eligibility criteria), with reasons	30	0.26-0.34	0.75
Participant(s)	4a. Diagnosis/disorder, diagnostic criteria, co-	72	0.67-0.76	0.74

	morbid conditions and concurrent therapies. In addition for series: Eligibility criteria for participants			
	4b. Settings and locations where the data were collected	45	0.40-0.50	0.83
Interventions	5. The interventions for each period with sufficient details to allow replication, including how and when they were actually administered	87	0.84-0.91	0.73
Outcomes	6a. Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	61	0.56-0.65	0.90
	6b. Description and measurement properties (validity and reliability) of outcome assessment tools	75	0.71-0.80	0.80
	6c. Any changes to trial outcomes after the trial commenced, with reasons	38	0.34-0.43	0.80
Sample size	7a. How sample size was determined	87	0.84-0.91	0.64
	7b. When applicable, explanation of any interim analyses and stopping guidelines	40	0.36-0.45	0.76
	Random Assignment			
Sequence generation	8a. Whether the order of treatment periods was randomized and method used to generate allocation sequence	79	0.75-0.83	0.85
	8b. When applicable, type of randomization; details of any restrictions (e.g. pairs, blocking)	29	0.25-0.34	0.87
	8c. Full, intended sequence of periods	33	0.28-0.37	0.83
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	23	0.19-0.27	0.92

Implementation	10. Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	16	0.13-0.20	0.84
Blinding	11a. If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	58	0.53-0.63	0.90
	11b. If relevant, description of the similarity of interventions	41	0.36-0.46	0.78
Statistical methods	12a. Statistical methods used to compare interventions for primary and secondary outcomes	71	0.66-0.75	0.73
	12b**. 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)	37	0.32-0.41	0.73
	12c. Statistical methods used to check assumptions (e.g. carry-over effect, period effects, intra-subject correlation, etc.)	29	0.25-0.34	0.85
Ethics	13. Whether the report represents a research study and if so, whether institutional ethics approval was sought	39	0.35-0.44	0.88
	Results			
Participant flow (a diagram is strongly recommended)	14a. Number and sequence of periods completed.	93	0.91-0.95	0.64
	14b. For series only: The number of participants who were enrolled, assigned to interventions, and analyzed for the primary	51	0.47-0.56	0.75

	outcome			
	14c. **For series only**: losses or exclusion of participants after treatment assignment, with reasons, and period in which this occurred, if applicable	38	0.34-0.43	0.88
Recruitment	15a. Dates defining the periods of recruitment and follow-up	78	0.74-0.82	0.72
	15b. Whether any periods were stopped early and/or whether trial was stopped early, with reason(s).	37	0.33-0.42	0.90
Baseline data	16. A table showing baseline demographic and clinical characteristics for each group	36	0.31-0.40	1.00
Numbers analyzed	17. For each intervention, number of periods analyzed. In addition for series: if quantitative synthesis was performed, number of trials included.	84	0.80-0.87	0.72
Outcomes and estimation	18. For each primary and secondary outcome, results for each period, the estimated effect size and its precision (e.g. 95 confidence interval); a visual depiction is recommended. In addition for series: if quantitative synthesis was performed, group estimates for each period or intervention	32	0.27-0.36	0.93
Ancillary analyses	19a. For binary outcomes, presentation of both absolute and relative effect sizes is recommended	32	0.27-0.36	0.75
	19b. Results of any other analyses performed, including assessment of carry-over effects, period effects, intra-subject correlation. In addition for series: If done, results of sub-group analyses	37	0.35-0.45	0.71
Harms	20. All harms or unintended effects for each intervention (for specific guidance see CONSORT for harms)	60	0.55-0.64	0.74

	Discussion			
Limitations	21. Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	72	0.67-0.76	0.68
Generalizability	22. Generalizability (external validity, applicability) of the trial findings	54	0.69-0.78	0.69
Interpretation	23. Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	61	0.66-0.75	0.73
	Other information			
Registration	24. Registration number and name of trial registry	3	0.02-0.06	0.71
Protocol	25. Where the full trial protocol can be accessed, if available	34	0.29-0.38	0.87
Funding	26. Sources of funding and other support (such as supply of drugs), role of funders	36	0.31-0.40	0.86

***** See the CENT Explanation and Elaboration paper for specific abstract guidance)

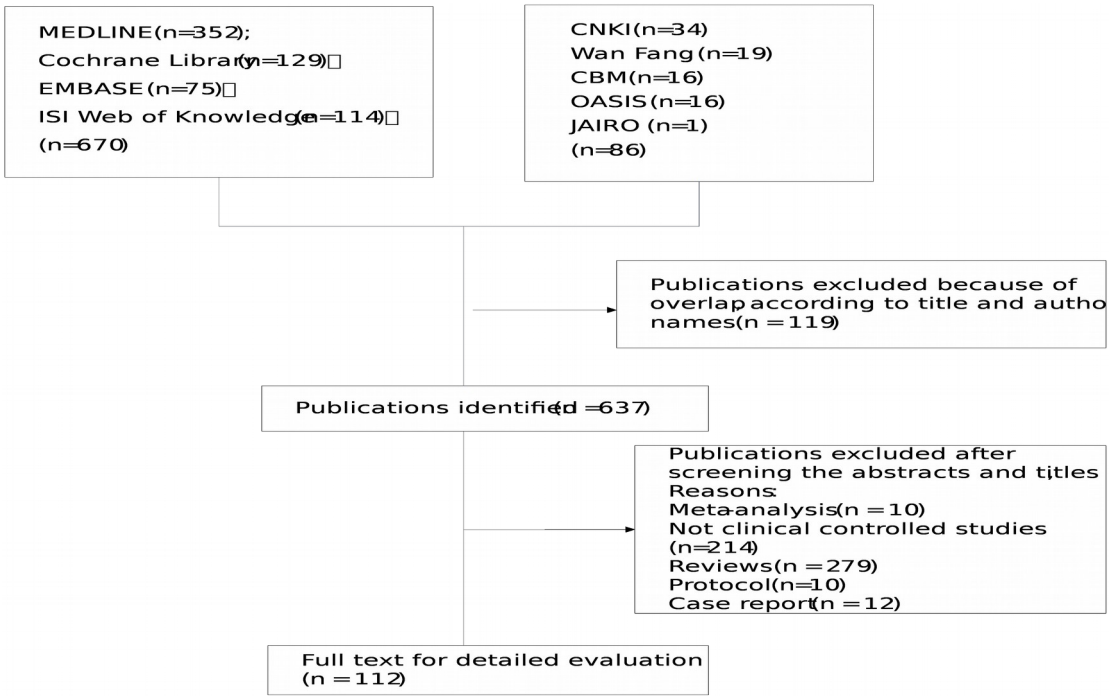
Table 3. Overall Reporting Quality Scores for 112 N-of-1 Trials, by Subgroup

Subgroup	Mean (SD) Score*	*P*
Impact factor of journal (n)		
Missing (n=18)	20.2 (8.3)	
Below the median value (n=40)	24.4 (8.0)	
Above the median value (n=54)	26.3 (8.9)	
Total Journals (N=112)		0.004
Year of Publication (n)		
1985-1989 (n=1)		
1990-1999 (n=42)	22.2 (7.9)	
2000-2009 (n=45)	28.5 (8.2)	
2010-2012 (n=24)	21.7 (8.8)	
Total reports (N=112)		0.001
First author's region (n)		
North America (n=52)	23.6 (8.9)	
Europe (n=33)	23.7 (8.7)	
Australia (n=19)	27.9 (8.2)	
Asia (n=7)	27.1 (8.3)	
South America (n=1)		
Total reports (N=112)		0.06

* A perfect score is 40 for single trials and 43 for combined trials

Captions

Figure. Summary of the article selection process in a systematic review of the quality of reporting of N-of-1 trials.


MEDLINE(n=352);
Cochrane Library(n=129);
EMBASE(n=75);
ISI Web of Knowledge(n=114);
(n=670)
CNKI(n=34)
Wan Fang(n=19)
CBM(n=16)
OASIS(n=16)
JAIRO (n=1)
(n=86)
Publications excluded because of overlap according to title and autho names(n =119)
Publications identified (n =637)
Publications excluded after screening the abstracts and titles
Reasons:
Meta-analysis(n =10)
Not clinical controlled studies (n=214)
Reviews(n =279)
Protocol(n=10)
Case report(n =12)
Full text for detailed evaluation (n =112)

Appendix

The search strategy:

((((((n-of-1) OR n-of-1 trial) OR single patient trial) OR single patient research) OR n-of-1 design) OR n-of-1 randomized controlled trial) AND ((crossover) AND randomization)