

A Proposal for a Simple and Useful Research Design for Evaluating the Efficacy of Acupuncture: Multiple, Randomized n-of-1 Trials

KAWAKITA Kenji¹⁾, SUZUKI Masao²⁾
NAMURA Kenji²⁾, TANZAWA Shouhachi¹⁾

1) Research Department of Japan Society of Acupuncture and Moxibustion

2) Department of Internal Medicine, Meiji University of Oriental Medicine

Abstract

[Aim] To develop an experimental design suitable for clinical acupuncture research.

[Design] Long-term n-of-1 trials (B-A-B-A design)

[Setting] University Hospital of Meiji University of Oriental Medicine

[Patient] A chronic bronchial asthma patient

[Intervention] Weekly acupuncture treatments for 10 min with de-qi were given. The initial 10 treatments (period B 1) were followed by 9 weeks baseline (A 1), a second period of 12 treatments (B 2), and a further baseline period (A 2).

[Main outcome measure] Asthma symptom score by diary

[Results] The patient's symptoms were clearly reduced during the treatment period but returned during the baseline period. These changes in asthma score were highly reproducible in this patient.

[Conclusion] The value of long-term n-of-1 trials in acupuncture research was clearly demonstrated, and it is suggested that the n-of-1 trial enables demonstration of the mi-byo-chi of acupuncture treatment. To increase the external validity of n-of-1 data, multiple, randomized n-of-1 trials are proposed as an appropriate design for clinical research into acupuncture.

Key words: acupuncture, AB design, n-of-1 trial, mi-byo-chi, clinical trial, asthma, multiple randomized n-of-1 trials

I. Introduction

It is well recognized that the randomized controlled trial (RCT) is the most powerful experimental design for generating strong evidence¹⁾. However, using the RCT to evaluate the clinical usefulness of acupuncture raises various issues that need to be resolved²⁾. One of the major problems is that the acupuncture treatment procedure is not fixed according to the disease or the patient's condition. Acupuncturists carefully select points for needle insertion that are individualized for each patient. This traditional approach to acupuncture treatment is very popular and has spread widely. The majority of

acupuncturists who are clinically well trained reject the use of fixed points or a predetermined set of points when treating patients, as they believe that acupuncture with an incorrect choice of points or inadequate procedures is ineffective. Their belief is usually based on clinical experience, not on evidence. They need to provide evidence to support their concept that the selection of points and other aspects of the process of acupuncture must be individualized in order to be effective. To evaluate the efficacy of acupuncture, various designs and types of control can be used, depending on the research question of the investigator³⁾.

II. N-of-1 trial as a useful design for acupuncture research

In the WHO guidelines on clinical research on acupuncture, single subject experimental designs (single case design, or n-of-1 trial) are introduced⁴⁾. N-of-1 trials (this term will be adopted in this paper) developed in the field of psychology, and have recently been adapted for clinical research⁵⁻⁹⁾. The statistical issues concerning the evaluation of the results have been clarified^{10,11)}.

The simplest design of an n-of-1 trial is a reversal design. Baseline data are collected repeatedly during period "A" and their stability is confirmed, without treatment. Then a specific intervention is applied during period "B". The changes in outcome data are evaluated by visual inspection of a graphical figure or by the usual non-parametric test for two groups¹²⁾. Repeating the two stages of the trial (A-B-A-B-A-B-) strengthens the plausibility of the results. The order BA instead of AB can be used when treatment is required urgently before the baseline period.

N-of-1 trials can evaluate the effectiveness of various specialized interventions in a number of patients who differ in several ways. They are easy to adopt for an exploratory study. The characteristics of the n-of-1 trial seem to be suitable for acupuncture research and the use of n-of-1 trials in acupuncture has been recommended^{4,13)}. However, the n-of-1 trial is not appropriate

in cases where acupuncture treatments have long-lasting or irreversible effects. Moreover it has been pointed out that the results of n-of-1 trials cannot be easily generalized.¹⁴⁾ Here we propose a unique protocol of n-of-1 trials that allows generalization from the results obtained from each patient attending an acupuncture clinic.

III. Long-term n-of-1 trials: a research design applied in an acupuncture clinic

In general, the majority of patients at acupuncture clinics seem to be regular attenders who visit the clinic each time their chronic illness deteriorates. Their complaints are treated successfully by acupuncture but will reappear after several weeks, months or years. Based on such a course of acupuncture treatment over time, we propose a new design for clinical research in acupuncture.

Figure 1 shows a hypothetical illustration of a long-term n-of-1 trial (ABAB design). The upper figure shows the severity of symptoms and the lower bar shows the baseline (A 1, 2) and intervention (B 1, 2) phases. In cases where patients' complaints are severe, the active intervention can be used first (BABA design). Another n-of-1 experimental design such as alternation may also be applicable. If the symptomatic changes produced by the intervention are very long lasting or permanent, a simple group comparison design should be used.

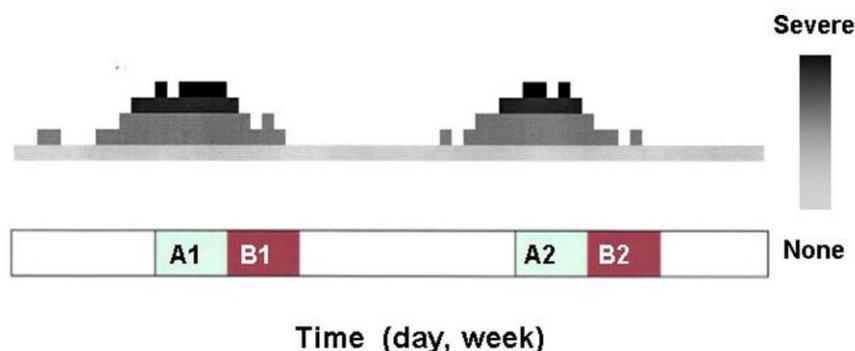


Figure 1. Hypothetical example of a long-term n-of-1 trial A 1, A 2: baseline, B 1, B 2: intervention Criteria for the onset of baseline data measurement should be determined in the protocol.

IV. An example of a long-term n-of-1 trial in an asthma patient

The effects of acupuncture on chronic bronchial asthma were examined by n-of-1 trials (BABA design) in one patient. The patient, who was receiving care from a medical doctor but was resistant to steroid treatment (oral and by inhalation), was recruited to the study. The patient received acupuncture treatment (once a week, 10 times, repeated for a second course). Acupuncture needles (0.16 mm in diameter, 40 mm in length) were inserted and retained for 10 minutes at the following meridian points bilaterally: LI 1, CV 12, LI 5, CV 4, and B 13. The severity of asthma was recorded by a diary of asthma symptoms, a VAS (visual analogue scale) of dyspnea, and Hugh-Jones classification. During the experiment the patient continued to receive steroids regularly.

Figure 2 clearly shows that every symptom measure gradually improved, almost completely disappearing after the initial 10 weekly acupuncture treatments, and then rapidly returned to the pre-treatment level 9 weeks after cessation of treatment (initial BA session). The second treatment course of 12 weeks produced more rapid and sustained improvement during the treatment, but the symptoms again returned after treatment had stopped. Changes in the measures before and after the second treatment were as follows: Symptomatic scores:

66 to 0, VAS of dyspnea: 87 to 0, H-J classification: IV to I.

These results show that a long-term n-of-1 trial may be useful for demonstrating the effects of acupuncture on patients over a long treatment period. This kind of situation, with repeated treatments for chronic conditions, may be very common in acupuncture clinics. So, we propose a unique protocol to allow generalization from the results obtained from long-term n-of-1 trials.

V. N-of-1 RCT (randomized controlled trial)

The clinical usefulness of n-of-1 trial has become widely recognized, but, in respect of evidence based medicine (EBM) overall, its lack of external validity reduces the strength of evidence that it can contribute. In a recent EBM textbook, the n-of-1 RCT was ranked as the strongest evidence for making treatment decisions¹⁵⁾. This high ranking of the n-of-1 RCT is mainly based on its high internal validity, that is, the n-of-1 RCT can make it possible to decide whether an intervention is suitable for a particular subject.

The simplest n-of-1 RCT is as follows: the patient is randomly allocated to two periods of interventions, either A/B or B/A. The efficacies of interventions A and B are evaluated by the use of appropriate outcome measures, and these alternating interventions continue until a significant difference is detected between their

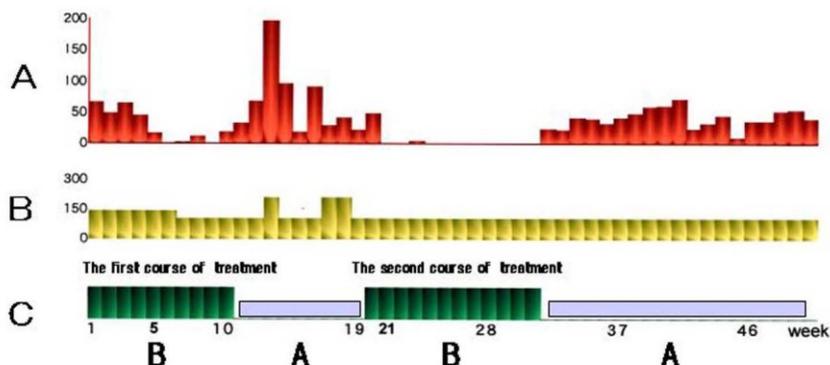


Figure 2. Example of a long-term n-of-1 trial of acupuncture for an asthma patient

A: score of the symptoms from asthma diary.

B: dose of drug used,

C: periods of treatment with acupuncture (B) and non-treatment baseline (A)

effects. If intervention A is superior to B, then A will be selected as better treatment for the subject.

Regarding the analysis of n-of-1 data, various methodological issues have been identified. Time series analysis was strongly recommended instead of conventional group comparison tests^{6,10}. Other statistical tests such as C-statistics have also been proposed as an indicator for an n-of-1 trial¹¹. Recent developments in computer technology make it possible to use the randomization test to analyze the data from n-of-1 trials¹⁶.

From the viewpoint of patient-oriented medicine, the n-of-1 RCT design is valuable and highly recommended. However, it should be noted that an n-of-1 RCT does not provide external validity. In Sackett's standard textbook of EBM, the n-of-1 RCT is not included in his classification of clinical trials and list of recommendations (Table 2), but he noted the importance of the design and stated guidelines for limitations on its application¹⁷. Every researcher agrees that a systematic review of homogeneous RCTs is the best EBM methodology for providing external validity.

We now propose a method to increase the external validity of n-of-1 trials by adding a randomization procedure in the group comparison.

VI. Multiple, randomized n-of-1 trials

We propose that multiple, randomized n-of-1 trials are a suitable design for increasing the external validity of a single n-of-1 study. Figure 3 shows the outline of the protocol. Patients who match the entry criteria are registered and randomly allocated into the acupuncture and control groups. Their condition or symptoms are treated by various acupuncture techniques, depending on the practitioner's method of diagnosis and treatment, the details of which should be reported in detail following the STRICTA (standards for reporting interventions in controlled trials of acupuncture) recommendations¹⁸. The effect on each patient is evaluated by a suitable statistical method such as a non-parametric test¹², and then the incidence of positive and negative results is compared between the two groups using a chi-square test.

Table 1. A hierarchy of strength of evidence for treatment decisions

1	N-of-1 RCT
2	SR of randomized trials
3	Single randomized trial
4	SR of observational studies addressing patient-important outcomes
5	Single observational study addressing patient-important outcomes
6	Physiologic studies (studies of blood pressure, cardiac output, exercise capacity, bone density, and so forth)
7	Unsystematic clinical observations

Table 2. Levels of evidence and grades of recommendations

Grade of recommendation	Level of evidence	Designs of clinical trials
A	1 a	SR (with homogeneity) of RCT
	1 b	Individual RCT (with narrow confidence interval)
	1 c	All or none
B	2 a	SR (with homogeneity) of cohort studies
	2 b	Individual cohort study (including low-quality RCT)
	2 c	"Outcome" research
C	3 a	SR (with homogeneity) of case-control study
	3 b	Individual case-control study
D	4	Case series (and poor-quality cohort and case control studies)
D	1 a	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

SR: systematic review

RCT: randomized controlled trial

To conduct this protocol successfully, several issues should be considered. The symptom should be stable over a long period and responsive to the intervention. The severity of the major symptom or the overall condition should be recorded daily during the experimental period by simple questionnaire or VAS scale. When the symptom appears to be stable (an essential inclusion criterion), baseline data are collected (period A: days, weeks or months), then the intervention is applied repeatedly (period B: days, weeks or months). Follow-up data are also collected. The interventions should be repeated at least twice to increase the reliability of results. This protocol is easy to conduct if suitable patients can be recruited. If the sample size is large enough to allow a subgroup analysis, the effectiveness of various combinations of symptoms and methods of acupuncture treatment may also be examined by the incidence of positive or negative results.

VII. The concept of "mi-byo-chi " for the acupuncture treatment

In general, the majority of patients at acupuncture clinics are regular attenders who visit to the clinic when they feel that their symptoms are getting worse, in order to restore their health. If the treatment is performed when the symptoms are not too severe, the results will be better than those obtained when the symptoms are becoming more severe. Figure 4 schematically illustrates the concept of "mi-byo-chi ".

The borderline between health and disease is not completely clear. In the ancient Chinese literature (the Yellow Emperor's textbook), the concept of mi-byo-chi was introduced. The "Mi-byo" means that the condition is pre-symptomatic, and "chi" means treatment, so the phrase indicates the importance of giving treatment before the symptoms become severe. When the treatment

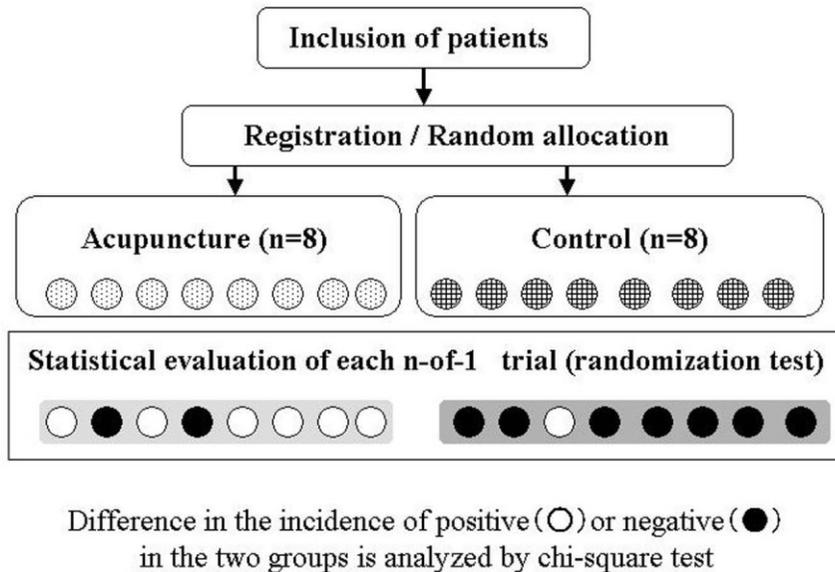


Figure 3. A block diagram of the protocol of multiple randomized n-of-1 trials
 Patients are randomly allocated to each group then the effect of acupuncture or control intervention in each patient is evaluated by an n-of-1 trial (the randomization test can be used). Incidences of positive or negative results are analyzed statistically. (using chi-square test). In this case, chi square=6.3349 and p=0.0117. However, small samples require Yates' continuity correction, with which the results become, chi-square=4.063, p=0.0438. This result indicates the external validity of n-of-1 trials.

Concept of “未病治 (mi-byo-chi)”

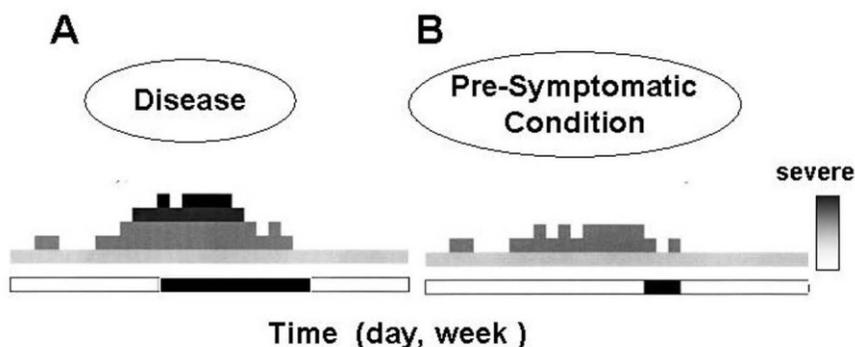


Figure 4. A schematic illustration of the concept of the "mi-byo-chi"
 A: treatment of the developed disease takes a long time,
 B: treatment of the pre-symptomatic condition is rapidly effective.

(thick black band) is applied to a condition that is less severe (B), the symptoms are abolished more rapidly than when it is applied to a condition that is more severe (A) (Fig. 4).

This concept clearly highlights the importance of the preventive aspect of acupuncture treatment. For evaluating the validity of the concept of "mi-byo-chi", the proposed multiple randomized n-of-1 trial may be applicable and would be worth conducting in a large sample in order to increase both internal and external validity of the clinical trial and provide stronger evidence.

Acknowledgement

The authors wish to express their thanks to the late Dr. Kuwata for his valuable suggestions and discussions. This study was supported by the Foundation for Training and Licensure Examination in Anma-Massage-Acupressure, Acupuncture and Moxibustion.

References

- 1) Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence based Medicine. 1 st ed. Churchill Livingstone. 1998.
- 2) White A, Filshie J, Cummings TM, Clinical trials of acupuncture: consensus recommendations for optimal treatment, sham controls and blinding. *Complement Ther Med* 2001; 9: 237-45.
- 3) Sherman KJ, Lao LX, MacPherson H et al. Matching acupuncture clinical study designs to research questions. *Clin Acupunct Orient Med*. 2001; 3: 12-5.
- 4) WHO Regional Office for the Western Pacific. Guidelines for clinical research on acupuncture. WHO Regional Publications. Western Pacific Series No.15. 1995.
- 5) Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial : Clinic-al usefulness. *Ann Intern Med*. 1990; 112: 293-9.
- 6) Vincent CA. The treatment of tension headache by acupuncture: a controlled single case design with time series analysis. *J Psychosom Res*. 1993; 34: 553-61.
- 7) Kawakita K, Okada K, Kuwata S. Application of n-of-1 trial and C statistics on acupuncture research, Proceeding of the workshop at WFAS in New York. 1996.

- 8) Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy-randomized trials in individual patients, *New Engl J Med.* 1986; 314: 889-92.
- 9) Mahon J, Laupacis A, Donner A, Wood T. Randomized study of n-of-1 trials versus standard practice. *BMJ.* 1996; 312: 1069-74.
- 10) Horne GP, Yang HCK, Ware WB. Time series analysis for single-subject designs. *Psychol Bull.* 1982; 91: 178-89.
- 11) Kawai I, Kawamoto H, Ohkouchi H. Applications of C statistic to test the treatment-effects in single-subject designs. *Jpn J Behav Analysis.* 1988; 2: 36-47.
- 12) Edgington ES. Nonparametric tests for single-case experiments. In: Kratochwill TR, Levin JR (eds). *Single-case research design and analysis.* London. Lawrence Erlbaum Associate. 1992: 133-57.
- 13) Kuwata S. Introduction of single subject designs as new experimental designs (II) Analytical evaluation of experimental data. *Jpn J Acupunct Mox.* 1993; 43: 36-43.
- 14) Kazdin AE. *Single-case research designs.* New York. Oxford University Press. 1982: 368.
- 15) Guyatt GH, Drummond R. *User's guided to the medical literature: Essentials of evidence-based clinical practice.* Am Med Assoc Press. 2002.
- 16) Todman JB, Dugard P. *Single-case and small-n experimental designs: A practical guide to randomization tests,* Lawrence Erlbaum Assoc Inc. 2000.
- 17) Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence based medicine. (2 nd ed.)* Edinburgh. Churchill Livingstone. 2000.
- 18) MacPherson H, White A, Cummings M et al. Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations. *Complement Ther Med.* 2001; 9: 246-9.