Single-Patient (N-of-1) Trials: A Pragmatic Clinical Decision Methodology for Patient-Centered Comparative Effectiveness Research

Naihua Duan, Ph.D.
Naihua.Duan@Columbia.Edu

Division of Biostatistics
Departments of Psychiatry
Columbia University

Outline

- Overview of Single patient (n-of-1) trials (SPT)
  - Clinical question
  - Basic protocol
  - Indications
  - Pros and cons
- IRB requirements
- Design of SPT
  - Blocked randomization or counter-balance
  - Blinding vs open label
  - Physical washout vs. “analytic washout”
  - Standardize or adapt/personalize?
- Statistical methods: borrow from strength
- Methodological development needs
- Explanatory trials vs. pragmatic trials
- Discussions

Audience Poll #1

- What is your primary role in the VA?
  - Student, trainee, or fellow
  - Clinician
  - Investigator/researcher
  - Statistician, IT/data manager
  - Policy-maker/manager/administrator
  - Other

Audience Poll #2

- What is your experience with SPT?
  - Have never heard of SPT
  - Have heard of SPT vaguely
  - Have some idea what SPT is about
  - Familiar with SPT protocol
  - Have considered conducting SPT
  - Have participated in SPT practice and/or investigations

Single Patient (N-of-1) Trials: Clinical Question

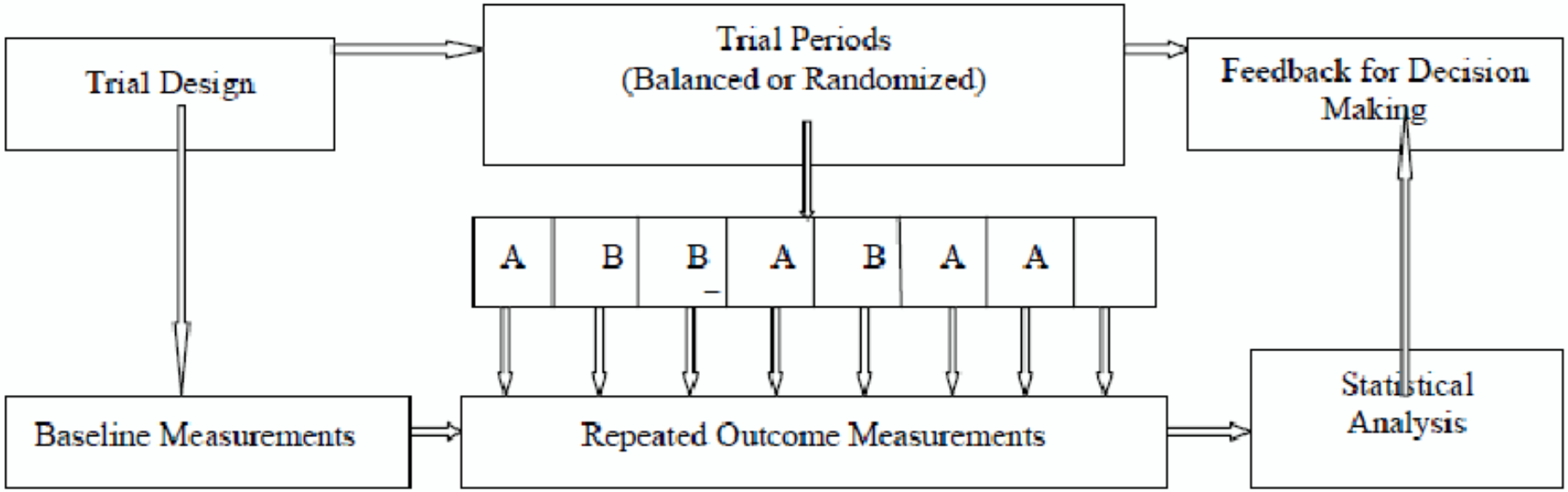
- Patient with chronic condition such as chronic pain
- Uncertainty Re: the comparative effectiveness of treatment options for this specific patient
  - Lack of existing research evidence
  - Potential for heterogeneity of treatment effects (Kravitz, Duan, Braslow 2004)
- Possible solution: conduct patient-centered comparative effectiveness investigation within this specific patient to inform his/her clinical decision

Single Patient (N-of-1) Trials: Basic Protocol

- Within patient multiple cross-over trials
- Assign time intervals (e.g., weeks) to alternate treatment options
- Collect outcome measures over time
- Compare outcomes under each treatment option
- Select treatment option with superior performance

Go to Figure 1 in PDF

Figure 1. Scheme for a prototypical SPT



Single Patient (N-of-1) Trials: Indications

- Duan, Kravitz, Schmid (2013), Table 1
- Chronicity and stability
  - On-going treatment for chronic conditions
  - Stable treatment effects
- Need for personalized knowledge
  - Lack of adequate evidence
  - Heterogeneity of treatment effects, one size might not fit all
- Quick effect onset and extinction
  - Quick onset of treatment effect
  - Negligible carry-over effect
  - No irreversible effects
- Examples: fibromyalgia, chronic pain, attention deficit-hyperactivity disorder, insomnia, asthma, chemotherapy-associated nausea and vomiting, and allergic rhinitis

Single Patient (N-of-1) Trials: Pros and Cons

- Consistent with routine clinical practice
  - Pull, not push
- Potential to improve outcomes for individual patients
  - Empirical evaluation is warranted
- Infrastructure needs
  - Application of mobile health (mHealth) technology
- Financial mechanism needs

PREEMPT Study

- Personalized Research for Monitoring Pain Treatment (PREEMPT)
- SPT using mHealth in Chronic Pain
- NINR-funded
- Infrastructure development (IT, Statistics)
- RCT to compare patients randomized to SPT vs. usual care

Literature (I)

- Duan N, Kravitz RL, Schmid CH. Single-Patient (N-of-1) Trials: A Pragmatic Clinical Decision Methodology for Patient-Centered Comparative Effectiveness Research. *Journal of Clinical Epidemiology*. 2013 Aug;66(8 Suppl):S21-8.
- Richard L. Kravitz and Naihua Duan (eds) and the DEcIDE Methods Center N-of-1 Guidance Panel (Naihua Duan, Ian Eslick, Nicole B. Gabler, Heather C. Kaplan, Richard L. Kravitz, Eric B. Larson, Wilson D. Pace, Christopher H. Schmid, Ida Sim, and Sunita Vohra). *Design and Implementation of N-of-1 Trials: A User's Guide*. AHRQ Publication No. 13-EHC122-EF. Rockville, MD: Agency for Healthcare Research and Quality; October 2013. Will be available online at effectivehealthcare.ahrq.gov.

Literature (II)

- Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Med Care*. 2011 Aug;49(8):761-8.
- Larson EB. N-of-1 trials: A new future? J Gen Intern Med. 2010 Sep;25(9):891-2.
- Zucker DR, Ruthazer R, Schmid CH. Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: methodologic considerations*. Journal of Clinical Epidemiology*, 2010. **63**(12): p. 1312-1323.
- Kravitz RL, Paterniti DA, Hay MC, Subramanian S, Dean DE, Weisner T, Vohra S, Duan N. Marketing therapeutic precision: Potential facilitators and barriers to adoption of n-of-1 trials. Contemp Clin Trials. 2009 Sep;30(5):436-45.
- Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What ever happened to N-of-1 trials? Insiders' perspectives and a look to the future. Milbank Q. 2008 Dec;86(4):533-55.
- Zucker DR, et al. Lessons learned combining N-of-1 trials to assess fibromyalgia therapies*. Journal of Rheumatology*, 2006. **33**(10): p. 2069-2077.
- Zucker DR, et al. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment*. Journal of Clinical Epidemiology*, 1997. **50**(4): p. 401-410.

Audience Poll #3

- Are you likely to participate in SPT practice and/or investigations in the next ten years?
  - Not at all
  - Somewhat possible
  - Maybe, 50/50
  - Likely
  - Certainly or almost certainly

IRB Requirements for SPT

- Human subjects research or quality improvement?
- Does intention to publish render a quality improvement program subject to regulations for human subjects research?

IRB Requirements for SPT: Does SPT Aim to Produce Generalizable Knowledge?

- Research means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge. (45 CFR 46.102(d))
- Primary objective for SPT is often to produce specific knowledge for individual patient, not to produce generalizable knowledge
  - Self-contained population for production and consumption of specific knowledge, not to export generalizable knowledge from trial population to consumption population
  - Generalizable knowledge might result from SPTs as a by-product, but not as the primary objective

IRB Requirements for SPT: SPT for QI

- Application of SPT for QI might not be subject to regulations for human subjects research
- "Protecting human subjects during research activities is critical and has been at the forefront of HHS activities for decades. In addition, HHS is committed to taking every appropriate opportunity to measure and improve the quality of care for patients. These two important goals typically do not intersect, since most quality improvement efforts are not research subject to the HHS protection of human subjects regulations." (http://answers.hhs.gov/ohrp/questions/7281)

IRB Requirements for SPT: Publications from SPT

- '…the intent to publish is an insufficient criterion for determining whether a quality improvement activity involves research… Planning to publish an account of a quality improvement project does not necessarily mean that the project fits the definition of research… (http://answers.hhs.gov/ohrp/questions/7286)

Design of SPT: Randomization or Counter-Balance

- Objective: balance between treatment conditions in terms of potential confounding factors such as time trend
  - AAABBB very bad
  - ABABAB not so good
- Blocked randomization vs. simple randomization
  - Small block size OK, no concerns about selection bias
- Counter-balance (maybe with restricted randomization) might achieve better balance
  - ABBA or BAAB provides better protection against linear time trend than ABAB or BABA

Design of SPT: Blinding or Open Label?

- Blinding (when feasible) is often important for parallel group trials that aim to produce generalizable knowledge for future patients
  - Expectancy among trial participants might not generalize to future patients
- Concern might not apply to SPT that aims to inform future treatment decision for the patient undergoing trial
  - Expectancy might persist from trial period into "consumption period"

Design of SPT: Physical Washout or "Analytic Washout"

- Washout period often inserted between treatment periods to eliminate/reduce carryover effect
- Does not address time required for onset of new treatment effect
- Might prolong transition between treatment period
- Problematic for comparative effectiveness investigations with active treatments
- "Analytic washout" models outcome trajectory, attempting to project long term treatment effect, without physical washout
  - Requires frequent outcome measurements, say, daily within weeklong treatment period

Design of SPT: Standardize or adapt/personalize?

- Incorporate user preference?
- Selection and weighting of outcomes
- Selection of criteria and format for reporting
- CER or PCOR

Statistical Methods: Borrow from Strength

- Individual's own SPT data most informative about his/her future treatment decisions
- Caveat: individual SPT usually of limited duration, with limited precision in estimated treatment effects
- Empirical Bayes methods can be used to "borrow from strength", combining index patient's own data with aggregate data from other patients with similar conditions, to provide more precise treatment effect estimates
  - Shrinkage estimator

Methodological Development Needs

- Analytic strategies to deal with onset and carryover effects
- Effective ways to summarize SPT findings for end-users (patients and their providers)
- Use of sequential stopping rules
- Use of responsive-adaptive designs with skewed randomization to incorporate partial information available

Explanatory vs. Pragmatic Trials

- Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis. 1967 Aug;20(8):637-48. Republished: J Clin Epidemiol. 2009 May;62(5):499-505. [Example]
- Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA. 2003 Sep 24;290(12):1624-32.

EXPLANATORY AND PRAGMATIC ATTITUDES IN THERAPEUTICAL TRIALS

DANIEL SCHWARTZ and JOSEPH LELLOUCH

Unité de Recherches Statistiques, Institut National de la Santé et de la Recherche Medicale, 94 Villejuif, France

(*Received* 6 *January* 1967; *in revised form* 24 *March* 1967)

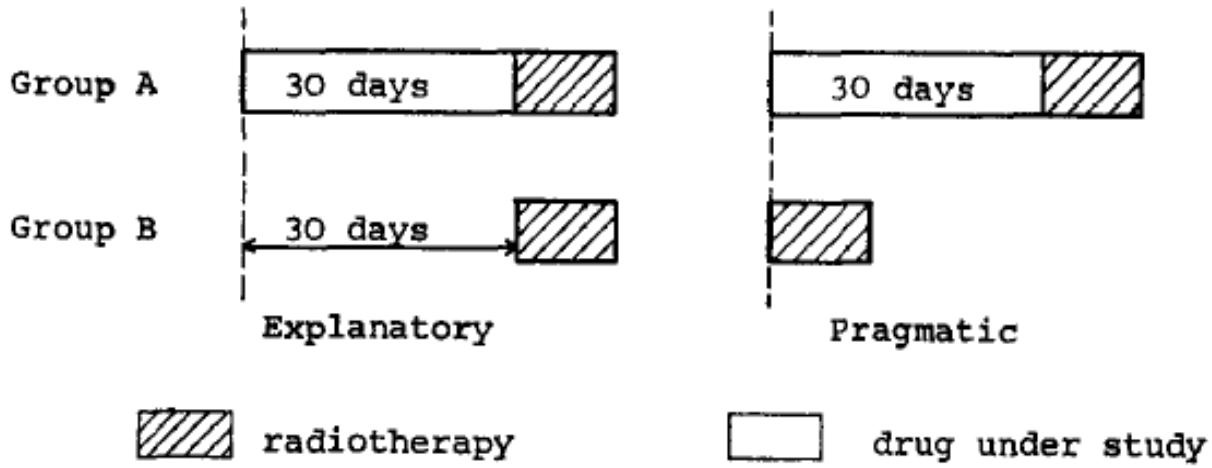
IT IS the thesis of this paper that most therapeutic trials are inadequately formulated, and this from the earliest stages of their conception. Their inadequacy is basic, in that the trials may be aimed at the solution of one or other of two radically different kinds of problem; the resulting ambiguity affects the definition of the treatments, the assessment of the results, the choice of subjects and the way in which the treatments are compared.

1.1. *"Equalized" or "optimal" conditions*

Consider a trial of anti-cancer treatments in which radiotherapy alone is to be compared with radiotherapy preceded by the administration of a drug which has no effect by itself but which may sensitise the patient to the effects of radiation. Suppose the drug is to be administered over a 30-day period. The "radiotherapy alone" group may then be handled in two different ways (Fig. 1):

(a) radiotherapy may be preceded by a blank period of 30 days, so that it is instituted at the same time in each group;

(b) radiotherapy may be instituted at once, thereby carrying it out at what is most probably the optimal time.



Neither procedure can be said to be "better" than the other. The first allows us to compare two groups which are alike from the radiotherapy point of view and which differ solely in the presence or absence of the drug. It therefore provides an assessment of the sensitising effect of the drug and gives valuable information at a biological level. The second procedure enables us to compare two treatments under the conditions in which they would be applied in practice. We distinguish the two procedures as stemming from two different approaches to the trial, the first *explanatory*, the second *pragmatic*.

In this example, the two approaches arise out of the complex nature of the treatments. When two treatments each consist of a series of components of which one is particularly to be studied, the other components may be carried out under either "equalized" or "optimal" conditions. The first possibility provides information on the effects of the key component, while the second compares two complex treatments as a whole under practical conditions.

Distinct Research Questions Neither Right Or Wrong

- Explanatory: lab condition, control for contextual factors – “everything being equal…”
  - Inform development of new treatments
- Pragmatic/practical: naturalistic condition, incorporate contextual factors – treatment bundle
  - Inform clinical and policy decision-making
  - Usual focus for SPT
- Scientific research vs. QI/engineering investigation

Discussions

- Single-patient (n-of-1) trials can be a useful tool for treatment decisions for on-going treatment for chronic conditions consistent with indications discussed in Duan, Kravitz, and Schmid (2013), Table 1.
- Effectiveness of SPTs in improving long term patient outcomes needs to be established empirically in studies such as PREEMPT.
- Broad implementation of SPTs requires solution of infrastructure needs and implementation issues.

Audience Poll #4

- Are you likely to participate in SPT practice and/or investigations in the next ten years?
  - Not at all
  - Somewhat possible
  - Maybe, 50/50
  - Likely
  - Certainly or almost certainly

Thank you!

- Questions?

- Naihua Duan, Ph.D.
  Naihua.Duan@Columbia.Edu

Is There an Attention Bias? (I)

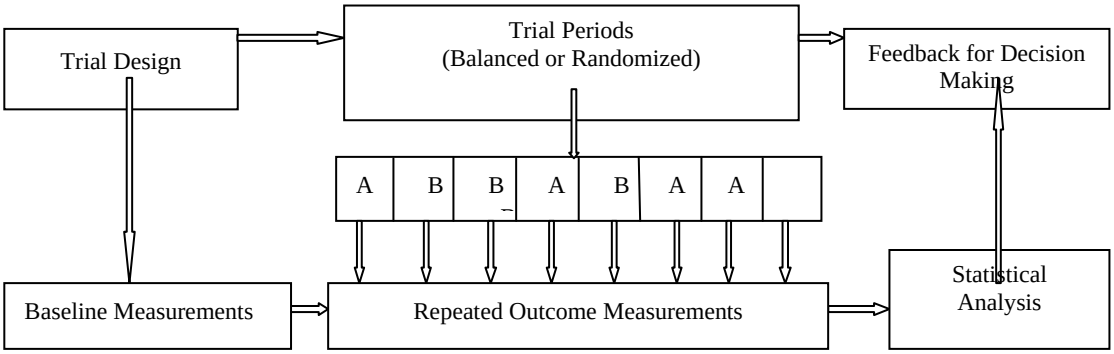
- Medication vs. Therapy
  - More intensive contact with clinician among therapy patients
- Explanatory trial
  - Difference in attention intensity considered a confounding factor
    - There is an attention bias
  - Control/equalize all contextual factors
    - Impose non-therapeutic clinician contact among medication patients?

Is There an Attention Bias? (II)

- Pragmatic trial
  - Difference in attention intensity is natural for the way each treatment is delivered
  - Compare treatment bundles, not isolated treatment ingredients
    - "Medication + low attention" vs.

      "Therapy + high attention"
  - There is no attention bias
  - Incorporate/optimize all contextual factors

Figure 1. Scheme for a prototypical SPT



ELSEVIER

Journal of Clinical Epidemiology 66 (2013) S21−S28

Journal of Clinical Epidemiology

Single-patient (n-of-1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research

Naihua Duan[a,*], Richard L. Kravitz[b], Christopher H. Schmid[c]

[a]Division of Biostatistics, Department of Psychiatry, Columbia University, 1051 Riverside Drive, Unit 48, New York, NY 10032, USA
[b]Division of General Medicine, Department of Internal Medicine, University of California, Davis, 4150 V. Street, Suite 2400 PSSB, Sacramento, CA 95817, USA
[c]Center for Evidence-Based Medicine, Department of Biostatistics, School of Public Health, Brown University, 121 S Main Street, Box G-S121-8, Providence, RI 02912, USA

Accepted 22 April 2013

Abstract

Objective: To raise awareness among clinicians and epidemiologists that single-patient (n-of-1) trials are potentially useful for informing personalized treatment decisions for patients with chronic conditions.

Study Design and Setting: We reviewed the clinical and statistical literature on methods and applications of single-patient trials and then critically evaluated the needs for further methodological developments.

Results: Existing literature reports application of 2,154 single-patient trials in 108 studies for diverse clinical conditions; various recent commentaries advocate for wider application of such trials in clinical decision making. Preliminary evidence from several recent pilot acceptability studies suggests that single-patient trials have the potential for widespread acceptance by patients and clinicians as an effective modality for increasing the therapeutic precision. Bayesian and adaptive statistical methods hold promise for increasing the informational yield of single-patient trials while reducing participant burden, but are not widely used. Personalized applications of single-patient trials can be enhanced through further development and application of methodologies on adaptive trial design, stopping rules, network meta-analysis, washout methods, and methods for communicating trial findings to patients and clinicians.

Conclusions: Single-patient trials may be poised to emerge as an important part of the methodological armamentarium for comparative effectiveness research and patient-centered outcomes research. By permitting direct estimation of individual treatment effects, they can facilitate finely graded individualized care, enhance therapeutic precision, improve patient outcomes, and reduce costs. © 2013 Elsevier Inc. All rights reserved.

Keywords: Adaptive trial; Bayesian method; Borrow from strength; Carryover effect; Crossover trial; Sequential trial; Washout

1. Introduction

A variety of critiques have been raised on the current paradigm for producing clinical knowledge. The core criticisms are that: (1) it is practically impossible to conduct standard parallel-group randomized controlled trials (RCTs) to address all clinically important questions, even those restricted to comparative effectiveness of drugs and devices [1,2]; (2) clinical evidence generated in those RCTs has poor generalizability and therefore limited applicability to real patients seen in ordinary practices [3,4]; and (3) treatments shown to be safe and effective on average may deliver an uneven mix of risks and benefits to individual patients, a problem known as heterogeneity of treatment effects (HTE; [4−10]). Furthermore, practitioners usually have little leverage over the choice of research topics or participation in the generation and interpretation of evidence. This represents an important missed opportunity to facilitate the development of a learning health care system ‘‘to generate and apply the best evidence for the collaborative health care choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care; and to ensure innovation, quality, safety, and value in health care’’ [11].

Single-patient trials, also known as n-of-1 trials and individual-patient trials, have the potential to address these critiques. Single-patient trials are multiple-period crossover experiments comparing two or more treatments within individual patients. Unlike parallel-group RCTs, single-patient trials can be used to estimate individual treatment effects directly. This allows single-patient trials to identify the best

Conflict of interest: The authors have no conflicts of interest related to the content of this article.

This study was presented at the Fourth DEcIDE Symposium on Comparative Effectiveness Research Methods: From Efficacy to Effectiveness; June 12−13, 2012; Rockville, MD.

* Corresponding author. Tel.: +212-543-5589.
E-mail address: Naihua.Duan@Columbia.Edu (N. Duan).

0895-4356/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jclinepi.2013.04.006

What is new?

- Single-patient (n-of-1) trials have struggled to gain acceptance among researchers, clinicians, and patients. The comparative effectiveness research and patient-centered outcomes research movements create an opportunity for such trials to serve as a clinical decision tool to inform personalized treatment decisions for patients with chronic conditions.
- Preliminary evidence from several recent pilot acceptability studies suggests that single-patient trials may have broader appeal among patients and clinicians than previously suspected. Bayesian and adaptive statistical methods hold promise for increasing the informational yield of single-patient trials while reducing participant burden, but are not widely used.
- Personalized applications of single-patient trials can be enhanced through further development and application of methodologies on adaptive trial design, stopping rules, network meta-analysis, washout methods, and methods for communicating trial findings to patients and clinicians.

treatment for each individual patient [8,12], thereby serving as a promising clinical decision tool for individual patients [13] in the spirit of patient-centered outcomes research (PCOR).

Gabler et al. [14] reviewed single-patient trials reported in the medical literature during 1985−2010, and identified 100 articles that reported on 108 studies enrolling a total of 2,154 patients. The studies addressed diverse clinical conditions, including neuropsychiatric (36%), musculoskeletal (21%), and pulmonary (13%). Examples of conditions to which single-patient trials have been applied successfully include fibromyalgia, chronic pain, attention-deficit hyperactivity disorder, insomnia, asthma, chemotherapy-associated nausea and vomiting, and allergic rhinitis.

The strengths and limitations for single-patient trials are reviewed in Table 1. To summarize, single-patient trials are suitable for evaluating long-term treatments for chronic conditions, with stable treatment response, quick onset of treatment effect, and modest or negligible carryover effects. The presence of HTE renders the evidence for individual treatment effects from single-patient trials especially informative [8]. Chronicity of the condition and stability of the treatment response provide opportunities for the initial investment in a single-patient trial to pay off through improvements in long-term patient outcomes. The lack of existing evidence creates a need for the evidence produced in single-patient trials. Slow-onset and/or carryover effects can compromise the validity of single-patient trials. Successful single-patient trials need to either ascertain that slow-onset and/or carryover effects are absent or negligible, or account for these transient effects with an appropriate washout period and/or an analytic strategy to untangle these effects from the true long-term treatment effect. At the same time, single-patient trials are not suitable for conditions that are acute or unrelentingly progressive (e.g., acute leukemia, where the clinician and patient might have only one chance to get it right); treatments causing permanent or only slowly reversible effects (e.g., surgery); and preventive treatments targeting conditions associated with uncommon, catastrophic outcomes (e.g., prevention of stroke in atrial fibrillation).

This article discusses how single-patient trials address the core critiques of evidence-based medicine (including its new guises, comparative effectiveness research [CER] and PCOR). We begin by reviewing the methodology of single-patient trials and their potential acceptability as a mainstream clinical decision tool. We then discuss applying Bayesian methods to combine single-patient trials across patients to produce estimates for individual treatment effects that are more stable than those generated by individual single-patient trials alone [15−18]. Finally, we discuss the methodological developments that could enhance the utility of single-patient trials.

2. Overview of methodology

Single-patient trials are multiple-period crossover trials conducted within individual patients to evaluate the comparative effectiveness of two or more treatments for each specific patient. As the methodology for single-patient trials has been discussed in detail elsewhere [13,15−23], we provide below only a brief overview. New directions for single-patient trial methodology are discussed in the ‘‘Methodological developments’’ section below.

We focus on the comparison of two treatments in the rest of this article; the extension to more than two treatments is relatively straightforward. Other types of single-subject designs, such as multiple baseline designs commonly used in social sciences research, differ in key features from single-patient trials and are not included in this discussion.

The unit of treatment assignment is a prespecified time period, say 1 week, during which the patient receives either treatment A or B. The duration of the treatment period is selected to allow each treatment an adequate time to manifest its effect. A washout period might be used between the two treatment periods to eliminate or reduce the carryover effect of the treatment used in the previous time period.

Treatment assignment is usually randomized and blocked to ensure good balance with respect to possible period effects. For example, within each block of two time periods, both treatments are used in one time period, so that the treatment assignments are randomized to either AB or BA in each block. The number of crossovers is usually prespecified, although some variation is often allowed to accommodate

Table 1. Indications and contraindications for single-patient trials (SPTs)

Feature	Description	Indication	Contraindication
Heterogeneity of treatment effects (HTE, [8])	Treatment effect varies across patients; one size does not fit all	With HTE, evidence based on specific patient is essential to personalize treatment decisions (e.g., serotonin reuptake inhibitors for treatment of depression)	Homogeneity of treatment effects (e.g., insulin [titrated to need] for reduction of blood glucose)
Chronicity	Long-term treatment for chronic condition	Chronicity allows knowledge gleaned from single-patient trials to inform future treatment decisions (e.g., gastroesophageal reflux disease)	Acute conditions (e.g., influenza) One-time treatment with long-lasting effects (e.g., surgery)
Stability	Stable treatment effect[a]	Stability ensures that knowledge gleaned from single-patient trials informs future treatment decisions	Lack of stability (e.g., in an individual whose dietary intake of vitamin K fluctuates widely over time, the effects of warfarin may be unstable relative to the effects of aspirin)
Effect onset and carryover	Transition periods between two treatment periods may be needed for the effect of previous treatment to extinguish, and the effect of new treatment to commence and stabilize. Insufficient length of either might confound estimation of long-term treatment effect	Negligible or modest duration for onset and carryover (e.g., short-acting psychostimulants for ADHD) allows single-patient trials to provide valid knowledge about long-term treatment effect, especially when accompanied with appropriate washout or analytic strategies to untangle slow onset and carryover effects from long-term treatment effect	Long duration of onset and/or carryover (e.g., long-acting medications)
Lack of adequate evidence	Existing clinical evidence not adequate to inform treatment decision for individual patients	Lack of adequate evidence creates the need for evidence to be gleaned from single-patient trials (e.g., effectiveness of prophylactic antibiotics in spinal cord injury patients with frequent urinary tract infections)	Adequate evidence: there is no need for further evidence from single-patient trials (e.g., effectiveness of HMG-CoA reductase inhibitors [statins] for reduction of cardiovascular risk in individuals with established coronary artery disease)

Abbreviation: ADHD, attention-deficit hyperactivity disorder.

[a] The assumption of stable treatment effect is weaker than the assumption of stable treatment outcome under both treatments. With the assumption of stable treatment effect, it is possible for treatment outcome to manifest a time trend, say, a gradual deterioration over time, as long as the trajectories are parallel for the two treatments, so that the difference between the treatments remains constant (stable). In other words, this assumption amounts to a requirement that treatment effect and time trend are additive, that is, there is no treatment $\times$ time interaction.

adverse events, in ways similar to the variations allowed in standard, parallel-group RCTs.

Clinical outcome is assessed repeatedly over time, at least once within each treatment period. The outcomes obtained during time periods with treatment A are compared with those obtained during treatment B to determine which treatment leads to preferred outcomes for the patient.

The analysis of single-patient trials includes informal evaluations such as visual inspection, simple statistical analysis such as the paired *t*-test, time series analysis to account for possible serial correlation, and Bayesian methods [13−18,20]. Among the studies reviewed by Gabler et al. [14], a quarter used visual inspection but no statistical comparison; among those that used statistical comparison, 59% used the *t*-test, 30% used nonparametric methods, 22% used regression modeling, 32% used pooled analysis, and only 7% used Bayesian methods (some used multiple methods).

On completion of the single-patient trial and the analysis, the patient and clinician meet to discuss the findings and decide on the treatment going forward. This decision must be based on a sound interpretation of the data obtained during the trial, including a careful assessment of the uncertainty involved that avoids overinterpreting effects that are driven by random noise in the data.

Single-patient trials primarily seek to produce specific clinical knowledge for each individual patient to inform the decision about which treatment to use in the long term. It is also possible to combine data across individual patients using Bayesian methods, to improve the clinical decision for individual patients beyond what can be accomplished using each patient's own data alone, and as a byproduct also to produce aggregate information that can be used to inform treatment decisions for other patients not participating in the trials [15−18]. Further discussions of Bayesian methods are given in the ''Bayesian methods'' section.

3. Acceptability

Several recent commentaries have advocated wider application of single-patient trials as a clinical decision tool [24−28]. Successful implementation, however, depends on the receptivity of stakeholders who control adoption of this methodology, especially clinicians and patients. We review in the following paragraphs several recent pilot acceptability studies that provide encouraging insights into the appeal of single-patient trials to these stakeholders.

Brookes et al. [29] conducted a pilot acceptability study among nine patients with osteoarthritis of the knee participating in two series of single-patient trials. One series ($n = 5$) compared a standard knee support with a heat-retaining support; the other series ($n = 4$) compared a nonsteroidal anti-inflammatory drug, diclofenac, with an analgesic, paracetamol. Patients were randomized to three pairs of crossovers, double blinded for the drug trials, and open label for the knee support trials (cannot blind). Daily diaries were taken to inform future treatment decisions. Qualitative, semistructured interviews were conducted with each patient at the start of the single-patient trial and again at the completion/termination to solicit patients' perspectives on trial participation, understanding of the protocol, initial expectations, and experience of participation. The authors concluded that the single-patient trial was an acceptable approach to individualize treatment decisions. Participants viewed the single-patient trial as a logical and accurate method that provided a fair opportunity for them to experience both treatments. They valued the personalized nature of the trial and its potential to offer improved treatment.

Kravitz et al. [30] conducted a pilot acceptability study among clinicians and patients to solicit potential facilitators and barriers to the adoption of single-patient trials. The study conducted phone interviews with 21 physicians (9 internal medicine, 5 family medicine, 4 rheumatology, and 3 pediatrics). They conducted focus groups, stratified by age, with 32 adult patients and parents of pediatric patients with at least one chronic condition. Despite initial lack of familiarity with the concept of the single-patient trial, both physicians and patients readily ''grasped the fundamental logic and appreciated the potential benefits'' of personalizing treatment decisions and enhancing doctor−patient relationship.

Nikles et al. [31] interviewed 12 Australian stakeholders recruited using purposive sampling through organizations considered likely to have an interest in single-patient trials to provide an optimally broad range of respondent categories (consumers, doctors, government, and industry). Stakeholders supported wider implementation of single-patient trials in a targeted fashion, with some caveats. They recognized the rationale behind single-patient trials to increase cost-effectiveness of government spending, to improve community health outcomes, and to maximize the effectiveness of individual treatment. They also recognized the potential benefits of single-patient trials, including targeting of therapy, reducing unnecessary prescribing, and reducing health care costs. Barriers recognized included constraints on doctors' time, doctors' acceptance, drug companies' acceptance, patient willingness, and cost. The authors identified several strategies for overcoming these barriers, namely (1) building single-patient trials into standard clinical consultation, (2) social marketing, (3) incentives such as payment for nurse practitioner support, (4) ensuring that the trials are suitable for real-world conditions and free for patients, (5) restricting the duration of the trials, and (6) obtaining funding for an independent organization to conduct the trials.

4. Bayesian methods

In most practical applications of single-patient trials, the number of crossovers that can be conducted is limited by resources and patient burden. Therefore, the comparative effectiveness estimated from the patient's own single-patient

trial data is often of limited statistical precision. It is therefore desirable to combine the index patient's own data with the data obtained from other patients who participated in similar single-patient trials to enhance the statistical precision available for the index patient. Bayesian methods, sometimes known as *borrowing from strength* [32], are useful for this purpose [15–18]. We present in the following paragraphs an overview of the Bayesian methods. Readers who are not interested in statistical details are welcome to skip to the ''Methodological developments'' section.

For a blocked design with two time periods in each block, the i-th patient's individual treatment effect (δ_i) can be estimated within each block as the difference in the observed outcome between the two treatments being compared. The block-specific individual treatment effect estimates can then be averaged across blocks within the i-th patient to provide a combined estimate for δ_i. The Bayesian model usually assumes δ_i to be drawn randomly from a normal distribution with mean δ_0 denoting the average treatment effect (ATE) for the population, and standard deviation τ denoting the variation in the individual treatment effects across patients (the HTE). The Bayesian framework requires placing prior distributions on these parameters that represent knowledge about these parameters before the study, say, from clinical knowledge or previous studies. When prior information is not available, noninformative prior distributions would be used. The ATE, δ_0, and the individual treatment effects, δ_i's, are then estimated using the posterior distribution for each parameter, usually using the posterior mean or median as the point estimate and the credibility interval based on the posterior distribution as the interval estimate.

The relationship between individual treatment effects and the ATE depends on the balance between the between-patient and within-patient variances [33]. When between-patient variance is small compared with the within-patient variance (i.e., little or no HTE), the individual treatment effects are very similar and close to the ATE. Alternatively, if between-patient variance is large compared with the within-patient variance (i.e., strong HTE), the individual treatment effects would be estimated to be close to the patient's observed treatment effect estimate with little or no borrowing from strength. In this situation, the strength (population information) to be borrowed from does not provide strong statistical information; therefore, the within-patient information dominates the between-patient information.

As a byproduct of the Bayesian method that aims to enhance clinical decisions for individual patients, the estimated ATE and HTE can also be used to inform treatment decisions for similar patients who did not participate in single-patient trials.

Some research studies can be designed either as a standard parallel-group RCT, or as an ensemble of single-patient trials. The parallel group design has more logistical simplicity, with no need to manage the crossovers and washout. On the other hand, single-patient trials might deliver greater power/precision than parallel group designs with the same number of patients, under the assumption that the carryover effect is either negligible or controlled for appropriately. With single-patient trials, the same patient serves as his/her own control; therefore, the idiosyncrasies unique to each patient are controlled for automatically. With parallel group designs, these idiosyncrasies contribute additional uncertainty, resulting in lower power/precision.

Related discussions on the precision for single-patient trials and traditional crossover trials are given in section 2.2.5 and Table 3 in the study of Zucker et al. [18]. They show that for a study with M patients and N paired-time periods, study precision is $M/(\tau^2 + 2\ \sigma^2/N)$, thus providing a way to calculate the tradeoff in sample size between the patients and time periods.

5. Methodological developments

As the single-patient trial receives wider attention in the CER/PCOR framework, further developments in its methodology, as discussed in the following paragraphs, can enhance the utility of single-patient trials in clinical practice.

Although a washout period between treatment periods is often used to guard against carryover effects, this strategy has limitations. Patients and clinicians might be dissatisfied with the withholding of active treatments during the washout, especially for comparative effectiveness single-patient trials with an active control. The lack of active treatments might also result in suboptimal patient outcomes during the single-patient trial. Furthermore, washout does not mitigate slow onset of the new treatment. Rather, the onset is deferred until after the end of the washout period, stretching the duration of the transient effects between the two treatment periods. Therefore, when designing a single-patient trial, there is a need to balance the benefits of washout (such as mitigating the potential bias of estimated treatment effects) against its limitations.

Hogben and Sim [5,6] used an innovative methodology to address the carryover effect without a washout period, taking daily measurements in each 3-day time period, but downweighting the earlier measures that are more susceptible to carryover effect. Zucker et al. [17] used a similar approach by analyzing only a single measurement at the end of each treatment period. Further analytic strategies can be used to model outcome trajectories during each treatment period, to untangle long-term treatment effects from transient effects owing to carryover of prior treatment and/or slow onset of new treatment. The ability for analytic strategies to deal with both carryover and slow onset is an important advantage over the usual strategy of a washout period. Further development and evaluation of these analytic methods are warranted.

Further investigation is warranted on effective ways to summarize findings from each single-patient trial for presentation to the patient and clinician. As there is diversity in the design and analysis for single-patient trials, there is also diversity in the decision process, in particular, how the findings

from the trial are presented. A variety of methods (such as visual inspection) used in the existing single-patient trial literature might be ineffective and vulnerable to overinterpreting the data when the observed treatment effect might be driven largely by random noise in the data.

As a patient-centered methodology, single-patient trials should prioritize incorporating user (patient and clinician) preferences, both in the design and analysis of the trial, and in the decision phase. A key consideration is the quantitative sophistication of the users. Some users might prefer to base their treatment decisions on the estimated treatment effect size and associated confidence interval for each outcome of interest. Others might prefer to know the posterior probability or odds for how the two treatments compare in terms of achieving prespecified goals, such as reducing the level of pain to a specific level, which takes uncertainty into consideration. However, such advanced feedback that recognizes the level of uncertainty in the information may be too difficult for many to grasp. Graphical presentations that show the trajectories of treatment responses over time for the two treatments are probably informative and comprehensible to all. Some patients may just prefer the pictures and may rely on their clinician to interpret the data. Thus, the single-patient trial needs to train patient and clinician to optimize the decision-making process while accommodating user preferences and maintaining scientific rigor.

Most single-patient trials are focused on prespecified treatments, such as two specific drugs to be compared. Although this highly structured design facilitates the implementation and interpretation of the trials (such as Bayesian methods for combining trial data across patients), a more flexible framework might be desirable in clinical applications, allowing each patient and their provider to choose the specific treatments of particular interest to them. With a flexible framework, it is more challenging to combine data across patients for the sake of borrowing from strength. A possible remedy is the use of network meta-analysis [34], to integrate direct (using trials that compared the specific treatments, say, A vs. B, directly) and indirect comparisons (using trials that compared A vs. B indirectly, say, through trials that compared A vs. C and trials that compared C vs. B), to maximize information available for the specific comparison of interest.

Essentially all single-patient trials are designed with a fixed 1:1 randomization ratio with a prespecified number of time periods. Alternatively, a response-adaptive design such as the play-the-winner design [35,36] could be used to allow the randomization ratio to adapt to the interim data obtained, so as to minimize patient exposure to the inferior treatment. The application of response-adaptive designs to single-patient trials calls for new methodological development in the multilevel framework (time periods nested within patients) to incorporate interim data from the index patient themselves and additional data from other similar patients. In applying a play-the-winner design, trial experience from other similar patients can be informative, but interim data from the index patient themselves should be weighted more heavily as being more directly relevant.

The common practice of prespecifying the number of crossovers might not always accommodate the needs of individual patients [20]. A more flexible strategy not prespecifying the number of crossovers can help accommodate the HTE across patients, for example, a sequential stopping rule [37] can be used to terminate the single-patient trial when a prespecified level of acceptable uncertainty has been reached. For patients who manifest dramatic differences between the two treatments, early stopping is indicated. For example, in a study by Guyatt et al. [19], a decision was made to terminate the trial after four time periods (two pairs of crossovers) because the unblinded data already convinced the patient and the clinician that one of the treatments was superior. Some patients might prefer more flexibility to allow for early termination when appropriate. Some might prefer the simplicity of a standard design with a fixed number of crossovers. Some patients' responses might fluctuate substantially, suggesting the need for more crossovers. Some patients might generate more consistent results and could benefit from early termination. Further research on adaptive trial design [38] and sequential stopping rules [37] for application to single-patient trials may facilitate individualization of the single-patient trial to enhance its scientific validity and user acceptability.

6. Summary and conclusions

After several decades of wandering in the wilderness [12], single-patient trials may be poised to emerge as an important part of the CER and PCOR methodological armamentarium. These trials render a number of benefits in clinical care. By permitting direct estimation of individual treatment effects, they can facilitate finely graded individualized care, enhance therapeutic precision, improve patient outcomes, and avoid unnecessary costs. New applications of Bayesian, adaptive, and sequential statistical methods hold promise for increasing the informational yield of single-patient trials while reducing participant burden.

Single-patient trials can deliver benefits that extend beyond patients participating in the trials. The aggregation of trial findings through Bayesian methods can inform treatment decisions for patients unaffiliated with the trials. In addition, by linking scientific methods to chronic disease management, single-patient trials encourage clinicians and patients to actively participate in real clinical learning communities. Supported by the right system-level infrastructure, clinicians can use single-patient trials to generate data that are not only scientifically valid but also immediately available and directly relevant to their patients. In this way, clinicians can obtain early, actionable feedback on their practices. It is therefore conceivable that single-patient trials might serve as a vehicle to facilitate the

transformation of clinical organizations into learning health care systems [11].

The way forward is not free of obstacles. Participation in single-patient trials requires time and effort, and although some patients and clinicians will be enthusiastic to participate, others will not be. The application of single-patient trials is restricted to chronic and symptomatic conditions with stable treatment outcomes. There is an ongoing tension among precision (longer trials are better), convenience (shorter are better), and clinical applicability (when a trial is too long, the results of earlier periods may no longer apply if the patient's condition has evolved).

Preliminary evidence from several recent pilot acceptability studies suggests that single-patient trials may be widely acceptable to patients and clinicians as an effective modality for increasing therapeutic precision. Because single-patient trials are akin to informal therapeutic trial-and-error procedures commonly used in clinical practice, clinicians may be amenable to upgrade their practice to incorporate this formal scientific procedure if the necessary infrastructure support is available. Even if just a small proportion of patients participate, the absolute number of eligible patients with chronic conditions who could benefit might be very large, making way for a prominent role for single-patient trials in clinical research and practice. If single-patient trials can succeed in luring only a fraction of eligible clinicians and patients into the scientific enterprise, the opportunity for transforming care through creation of high-functioning knowledge organizations is immense.

Acknowledgments

N.D.'s work on this article was supported in part by National Institutes of Health (NIH) Grant 7 P30 MH090322. R.L.K.'s work on this article was supported in part by NIH Grants 5 R01 MH079387, 5 R01 CA140419, 1 R01 NR013938, and UL1TR000002. The authors thank Ms. Kristen Greenlee for assistance with text processing.

References

[1] Brook RH. Possible outcomes of comparative effectiveness research. JAMA 2009;302:194−5.
[2] Tanenbaum SJ. Comparative effectiveness research: evidence-based medicine meets health care reform in the USA. J Eval Clin Pract 2009;15:976−84.
[3] Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282:1458−65.
[4] Greenfield S, Kravitz R, Duan N, Kaplan SH. Heterogeneity of treatment effects: implications for guidelines, payment, and quality assessment. Am J Med 2007;120:3−9.
[5] Hogben L, Sim M. The self-controlled and self-recorded clinical trial for low-grade morbidity. Br J Prev Soc Med 1953;7(4):163−79.
[6] Hogben L, Sim M. The self-controlled and self-recorded clinical trial for low-grade morbidity. Int J Epidemiol 2011;40:1438−54.
[7] Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. Trials 2010;11:85.
[8] Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q 2004;82(4):661−87.
[9] Longford NT. Selection bias and treatment heterogeneity in clinical trials. Stat Med 1999;18:1467−74.
[10] Tabery J. Commentary: Hogben vs the Tyranny of averages (Reprinted). Int J Epidemiol 2011;40:1454−8.
[11] Olsen LA, Aisner D, McGinnis JM, editors. The learning healthcare system: workshop summary. Washington, DC: National Academies Press; 2007.
[12] Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What Ever Happened to n-of-1 trials? Insiders' perspectives and a Look to the future. Milbank Q 2008;86(4):533−55.
[13] Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med 2011;8(2):161−73.
[14] 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.
[15] Nikles J, itchell 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.
[16] 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.
[17] 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.
[18] 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.
[19] 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.
[20] Guyatt G, Sackett D, Adachi J, Roberts R, Chong J, Rosenbloom D, et al. A clinician's guide for conducting randomized trials in individual patients. CMAJ 1988;139(6):497−503.
[21] Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial—clinical usefulness. Our 3-year experience. Ann Intern Med 1990;112:293−9.
[22] Larson EB. N-of-1 clinical-trials: a technique for improving medical therapeutics. West J Med 1990;152(1):52−6.
[23] Larson EB, Ellsworth AJ, Oas J. Randomized clinical-trials in single patients during a 2-year period. JAMA 1993;270:2708−12.
[24] Curtiss FR. Single-patient trial (SPT) method, substitute for expert opinion? J Manag Care Pharm 2002;8(6):519−20.
[25] Federman DG, Shelling ML, Kirsner RS. N-of-1 trials: not just for academics. J Gen Intern Med 2011;26:115.
[26] Larson EB. N-of-1 trials: a new future? J Gen Intern Med 2010;25:891−2.
[27] Rodnick JE. Australia: the N of 1 trial, an underappreciated research method. Fam Med 2006;38(1):63.
[28] Sackett DL. Clinician-trialist rounds: 4. why not do an N-of-1 RCT? Clin Trials 2011;8:350−2.
[29] Brookes ST, Biddle L, Paterson C, Woolhead G, Dieppe P. "Me's me and you's you": Exploring patients' perspectives of single patient (n-of-1) trials in the UK. Trials 2007;8:10.
[30] Kravitz RL, Paterniti DA, Hay MC, Subramanian S, Dean DE, Weisner T, et al. Marketing therapeutic precision: potential facilitators and barriers to adoption of n-of-1 trials. Contemp Clin Trials 2009;30:436−45.
[31] Nikles J, Mitchell GK, Clavarino A, Yelland MJ, Del Mar CB. Stakeholders' views on the routine use of n-of-1 trials to improve clinical care and to make resource allocation decisions for drug use. Aust Health Rev 2010;34(1):131−6.

[32] Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. Stat Med 1996;15:2733−49.
[33] Schmid CH, Brown EN. Bayesian hierarchical models. Methods Enzymol 2000;321:305−30.
[34] Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods 2012;3:80−97.
[35] Wei LJ, Durham S. Randomized play-winner rule in medical trials. J Am Stat Assoc 1978;73:840−3.
[36] Zelen M. Play winner rule and controlled clinical trial. J Am Stat Assoc 1969;64:131−46.
[37] Whitehead J. The design and analysis of sequential clinical trials. New York, NY: Wiley; 1997:328.
[38] Berry SM, Carlin BP, Lee JJ, Muller P. Bayesian adaptive methods for clinical trials. Boca Raton, FL: CRC Press; 2010.