BASS 2017
Savannah, GA

N-of-1 Trials: Do They Have a Role in Clinical Research?

Vernon M. Chinchilli, PhD
Department of Public Health Sciences
Penn State College of Medicine

Michele L. Shaffer, PhD
Department of Statistics
University of Washington

Introduction

- Physicians frequently use N-of-1 (single-patient) trial designs in an informal manner to identify an optimal treatment for an individual patient.

- Lillie et al (*Personalized Medicine* 2011) state the following: "N-of-1 or single subject clinical trials consider an individual patient as the sole unit of observation in a study investigating the efficacy or side-effect profiles of different interventions. The ultimate goal of an n-of-1 trial is to determine the optimal or best intervention for an individual patient using objective data-driven criteria. … n-of-1 trials demand serious attention among the health research and clinical care communities given the contemporary focus on individualized medicine."

Introduction

- N-of-1 clinical trials have been used quite frequently in education and behavioral science, but not very frequently in clinical research.

- An N-of-1 clinical trial that focuses exclusively on optimizing the primary outcome for a specific patient clearly may not be very useful for generalizability to a population of patients.

- A series or collection of N-of-1 clinical trials, however, could be generalizable.

Design Issues

- A typical N-of-1 clinical trial consists of a multi-way crossover design, such as ABAB, where A and B designate different treatments. Obviously, the physician-investigator needs to determine the appropriate number of treatments, the number of periods, and the length of each treatment period.

Period 1	Period 2	Period 3	Period 4
A	B	A	B

- Confounding factors can be minimized with the inclusion of more treatment periods, but more periods result in higher cost, a lengthier trial, and a higher risk of study withdrawal.

- Similar to crossover trials, N-of-1 clinical trials are suitable for chronic diseases but not for acute conditions.

Design Issues

- There is very low statistical power for the analysis of the primary outcome in an N-of-1 clinical trial. Therefore, prior to study onset, the physician-investigator needs to specify definitive criteria for treatment success.

- Randomization is an issue. If there only will be one patient, then it may be best to alternate the treatments, such as ABAB or ABABAB, and then randomize the treatments to the A and B designations.

Design Issues

- If there will be multiple N-of-1 clinical trials, then the physician-investigator can form sequences as in any crossover trial and randomize the patients to the sequences.

- There could be carryover effects from the treatments. Thus, the physician-investigator needs to determine whether placebo or active control wash-out periods would benefit study validity or risk patient safety.

- Blinding may not be possible for some of the interventions. Also, a patient may not wish to cross over to the next treatment period if the treatment during the current period is effective.

Design Issues

- Typically, physician-investigators assume minimal carryover effects in N-of-1 trials, but this may not be realistic.

- Some N-of-1 trials only capture information at the end of a treatment period in order to increase the independence of observations taken from the patient.

- Alternatively, if multiple observations are taken within a treatment period, then they can be weighted toward observations at the end of the period to minimize the impact of carryover.

Statistical Analysis

- The major feature of the statistical analysis in an N-of-1 clinical trial is the correlation among the repeated measurements from an individual patient. If there is only one N-of-1 trial with only a few measurements, then accounting for the correlation can be achieved via an autoregressive statistical model.

- If the time points are equally spaced, then a simple one-lag autoregressive model is as follows:

$$Y_j = \boldsymbol{X}_j^T \boldsymbol{\beta} + \varepsilon_j + \varphi \varepsilon_{j-1}, j = 1,2, \dots, p$$

where

Y_j is the response during the j^{th} period, $j = 1,2, \dots, p$

$\boldsymbol{X}_j$ is a vector of fixed-effects regressors

$\boldsymbol{\beta}$ is a vector of fixed-effects parameters

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p \; i.i.d. \sim N(0, \sigma^2), \varepsilon_0 = 0$

φ is an autoregressive parameter

Statistical Analysis

- This one-lag autoregressive model leads to the following variance and covariance expressions:

$$Var(Y_1) = \sigma^2,\ Var(Y_j) = (\varphi^2 + 1)\sigma^2, j = 2, \dots, p$$
$$Cov(Y_{j-1}, Y_j) = \varphi\sigma^2, j = 2, \dots, p$$
$$Corr(Y_{j-1}, Y_j) = \varphi/(\varphi^2 + 1), j = 2, \dots, p$$
$$Corr(Y_{j-2}, Y_j) = 0, j = 3, \dots, p$$

- Higher-lag autoregressive models are possible, but obviously there is a limit due to the small number of periods.

Statistical Analysis

- Example: Mahon et al (1999 *Chest*) conducted N-of-1 trials in patients with irreversible chronic airflow limitation. Patients alternated between periods of theophylline (T) and placebo (P). Table 4 in the article lists the data for symptom scores in seven patients:
  - five patients underwent a TPTPTP design
  - two patients underwent a TPTPTPTP design

- We applied the one-lag autoregressive model to the data for each of these seven patients. Two of the seven patients displayed statistically significant improvements in symptom scores from theophylline when compared to placebo.

Statistical Analysis

- Meta-analytic methods can be applied to combine the results from a series of N-of-1 trials.

- Zucker et al (*Journal of Clinical Epidemiology* 1997 and 2010) proposed Bayesian hierarchical models:
  - With few observations per patient and little information about within-patient variation, combined N-of-1 trials data may not support models that include complex variance structures.
  - Prior information with Bayesian models can be useful for increasing the precision of estimates but are very sensitive to prior assumptions about variance components.
  - Models with fixed treatment effects and common variances are robust and lead to conclusions that are similar to, though more precise than, single-period or single-crossover study designs.

Statistical Analysis

- A Bayesian model to consider for a continuous outcome in a series of N-of-1 trials for I individuals is as follows:

  $\boldsymbol{Y}_i = \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_i, i = 1,2, \dots, I$

  where

  $\boldsymbol{Y}_i$ is the $p_i \times 1$ vector of responses

  $\boldsymbol{\mu}_i$ is the $p_i \times 1$ vector of location parameters

  $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_I \; independent \sim N_{p_i}(\mathbf{0}, \boldsymbol{\Sigma}_i)$

  $\boldsymbol{\Sigma}_i$ is the $p_i \times p_i$ positive-definite variance-covariance matrix

- We assume prior probability distributions for $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ as

  $\boldsymbol{\mu}_i \sim N_{p_i}(\boldsymbol{\alpha}_i, \boldsymbol{\Psi}_i)$ and $\boldsymbol{\Sigma}_i \sim Wishart^{-1}_{p_i \times p_i}(\nu_i, \boldsymbol{\Psi}_i)$

  where

  $\boldsymbol{\alpha}_i$ is a $p_i \times 1$ vector of hyper-parameters

  $\boldsymbol{\Psi}_i$ is a $p_i \times p_i$ positive-definite matrix of hyper-parameters

  $\nu_i > p_i - 1$ is a degrees-of-freedom hyper-parameter

Statistical Analysis

- The marginal posterior probability distribution for $\boldsymbol{\mu}_i, i = 1,2,\ldots,I$, is a multivariate t, i.e.,

$$\boldsymbol{\mu}_i \,|\boldsymbol{Y}_i \sim t_{p_i}\left(\frac{\boldsymbol{\alpha}_i+\boldsymbol{Y}_i}{2},\ \frac{1}{\nu_i-p_i+2}\left\{\boldsymbol{\Psi}_i+\frac{1}{2}(\boldsymbol{Y}_i-\boldsymbol{\alpha}_i)(\boldsymbol{Y}_i-\boldsymbol{\alpha}_i)^T\right\}, \nu_i-p_i+2\right)$$

- As recommended by Zucker et al, we should borrow across the hyper-parameters of the I patients to render the analysis robust.

Statistical Analysis

- Schluter and Ware (*Statistics in Medicine* 2005) proposed a Bayesian approach for multiple N-of-1 trials when the response for each patient is the number of pairs of periods in which the experimental treatment out-performs the control treatment.

- Let X_i denote the response for the i^{th} patient, $i = 1,2, \ldots, I$. Schluter and Ware assumed that

  $$X_i \sim Binomial(n_i, \theta_i)$$

  where

  n_i denotes the number of pairs of periods for the i^{th} patient

  θ_i denotes the probability that the experimental treatment out-performs the control treatment within each pair of periods for the i^{th} patient

Statistical Analysis

- Schluter and Ware next assumed that $\theta_1, \theta_2, \ldots, \theta_I$ are independent and they imposed a two-stage prior probability distribution for each θ_i (instead of the usual beta prior probability distributional assumptions).

- This resulted in a posterior probability distribution for each θ_i that involves an integral equation, which is used to determine if $\Pr[\theta_i \geq 0.5]$ is large enough to claim that the experimental treatment is more effective than the control treatment.

CONSORT Statement

- Vohla et al (*BMJ* 2015) published a CONSORT statement for reporting the design, conduct, and analysis of N-of-1 trials. Important aspects to report on the study design include
  - Describe trial design, planned number of periods, and duration of each period (including run-in and wash out, if applicable)
  - *In addition for series:* Whether and how the design was individualized to each participant, and explain the series design
  - Whether the order of treatment periods was randomised, with rationale, and method used to generate allocation sequence
  - When applicable, type of randomisation; details of any restrictions (such as pairs, blocking)

CONSORT Statement

- Important aspects to report on the statistical analysis include
  - Methods used to summarize data and compare interventions for primary and secondary outcomes
  - *For series:* If done, methods of quantitative synthesis of individual trial data, including subgroup analyses, adjusted analyses, and how heterogeneity between participants was assessed
  - Statistical methods used to account for carryover effect, period effects, and intra-subject correlation

- Shamseer et al (*BMJ* 2015) published a follow-up article with more specific details about each of the CONSORT items.

The Future for N-of-1 Trials

- What does the future hold for N-of-1 trials?

- Obviously, N-of-1 trials could be useful as a small component of a Phase II research program, especially if dealing with a rare disease.

- Could N-of-1 trials be instrumental as Phase III clinical trials in the research program for the development of a new drug, biological product, or medical device?

The AsthmaNet BARD Trial

- AsthmaNet is a clinical trials network funded by the National Heart, Lung and Blood Institute (NHLBI) of the National Institutes of Health (NIH).

- An AsthmaNet trial that recently completed patient follow-up is entitled the “Best African American Response to Asthma Drugs (BARD)” trial.

- The www.ClinicalTrials.gov identifier is NCT01967173.

- The purpose of the BARD trial is to determine the best asthma treatment to add for Blacks who have asthma that is not well controlled on a low-dose inhaled steroid.

The AsthmaNet BARD Trial

- The BARD trial invokes a four-way crossover design. For convenience, the four treatment regimens are designated as A,B,C, and D for adolescents/adults in the following manner:
  - Run-in 1.0×ICS
  - Treatment A 1.0×ICS + LABA
  - Treatment B 2.5×ICS
  - Treatment C 2.5×ICS + LABA
  - Treatment D 5.0×ICS

  where
  - 1.0×ICS represents a low-dose inhaled corticosteroid, namely, 100 mcg fluticasone propionate BID
  - LABA represents a long-acting beta-agonist, namely, 50 mcg salmeterol BID

The AsthmaNet BARD Trial

- The BARD trial incorporated a design in which each participant received each of four treatment regimens over four 14-week periods (a four-way crossover design). For adults/adolescents, the four treatment sequences are as follows:

Sequence		Period 1	Period 2	Period 3	Period 4
#1		A	B	C	D
#2		B	D	A	C
#3		C	A	D	B
#4		D	C	B	A

- This crossover design has some optimal properties (uniform within sequences, uniform within periods, and balanced with respect to first-order carryover effects).

The AsthmaNet BARD Trial

- The BARD trial was tripled-blinded (patients, investigators, and biostatisticians).

- Randomization was stratified according to the nine clinical center partnerships.

- The BARD trial randomized
  - 294 adolescents/adults (12 years of age and older)
  - 280 children (5-11 years of age)

- Because the BARD trial did not include wash-out periods, the data from the first two weeks of each treatment period were not included in the data analysis.

The AsthmaNet BARD Trial

- The primary outcome in the BARD trial is the superiority of one treatment regimen compared to another treatment regimen using a variable based on a hierarchical determination from three asthma outcomes:
  - asthma exacerbations
  - annualized asthma control days (AACDs)
  - forced expiratory volume in one second (FEV_1)

- Step 1: If a BARD participant experiences fewer asthma exacerbations on one treatment regimen relative to another treatment regimen, then the treatment regimen that yields the fewer asthma exacerbations is deemed to be superior to the other treatment regimen and the process is terminated. If not, then continue to the next step.

The AsthmaNet BARD Trial

- Step 2: If a BARD participant experiences at least 31 fewer AACDs on one treatment regimen relative to another treatment regimen, then the treatment regimen that yields at least 31 more AACDs is deemed to be superior to the other treatment regimen and the process is terminated. If not, then continue to the next step.

- Step 3: If a BARD participant displays at least 5 percentage points higher in the % predicted FEV_1 at the end of the 14-week treatment regimen relative to another treatment regimen, then the treatment regimen that yields the higher % predicted FEV_1 is deemed to be superior to the other treatment regimen. If not, then the two treatment regimens are deemed to be "equivalent" or "tied" for that BARD participant.

The AsthmaNet BARD Trial

- For each BARD participant, this primary outcome variable is determined for all six pairwise comparisons of the treatment regimens:
  - A vs B
  - A vs C
  - A vs D
  - B vs C
  - B vs D
  - C vs D

- The AsthmaNet investigators did not define “absolute” responses, i.e., they did not define whether a BARD participant responds to any of the four treatment regimens. They only defined “comparative” responses.

The AsthmaNet BARD Trial

- In essence, the BARD trial is a large series of N-of-1 trials (294 adolescents/adults and 280 children) because it attempts to identify an optimal treatment regimen for each participant. In addition, it is possible to perform statistical inference at the population level.

- Also, notice that the BARD trial does not include replication of any of the four treatment regimens. This renders it very challenging to apply an autoregressive model to compare the four treatment regimens within an individual (N-of-1 trial).

The AsthmaNet BARD Trial

- For the A vs B comparison within the i^{th} BARD participant, $i = 1,2,\ldots,I$, define

$$Y_{i,AB} = \left\{ \begin{array}{r} +1,\ if \quad A > B \\ 0,\ if \quad A \approx B \\ -1,\ if \quad A < B \\ .\ ,\ \ if\ missing \end{array} \right\}$$

- A missing value occurs when a BARD participant withdraws from the trial at a point in time such that the A vs B comparison is not possible.

- Thus, there are six outcomes for the i^{th} BARD participant, $i = 1,2,\ldots,I$, namely, $Y_{i,AB}, Y_{i,AC}, Y_{i,AD}, Y_{i,BC}, Y_{i,BD}, Y_{i,CD}$.

The AsthmaNet BARD Trial

- Construct three logit functions for the A vs B comparison within the i^{th} BARD participant, $i = 1,2,\ldots,I$ as follows:

$$\mu_{i,AB(1)} = log_e\left\{\frac{\Pr[Y_{i,AB}=+1]}{\Pr[Y_{i,AB}=-1]}\right\}$$

$$\mu_{i,AB(2)} = log_e\left\{\frac{\Pr[Y_{i,AB}=+1\ or\ -1]}{\Pr[Y_{i,AB}=0]}\right\}$$

$$\mu_{i,AB(3)} = log_e\left\{\frac{\Pr[Y_{i,AB}=+1,0,\ or\ -1]}{\Pr[Y_{i,AB}=.]}\right\}$$

- The first logit represents the comparison of A superiority versus B superiority. The second logit represents the comparison of a superiority determination versus no superiority determination. The third logit represents the comparison of a non-missing observation versus a missing observation.

The AsthmaNet BARD Trial

- A generalized linear mixed-effects model for the A vs B comparison within the i^{th} BARD participant, $i = 1,2,\ldots,I$ is

$$\mu_{i,AB(1)} = \alpha_{AB(1)} + \left(\boldsymbol{X}_{i,AB} - \boldsymbol{X}_{Ref,AB}\right)^T \boldsymbol{\beta}_{AB(1)} + \left(\boldsymbol{Z}_{i,AB}\right)^T \boldsymbol{\gamma}_i$$

$$\mu_{i,AB(2)} = \alpha_{AB(2)} + \left(\boldsymbol{X}_{i,AB} - \boldsymbol{X}_{Ref,AB}\right)^T \boldsymbol{\beta}_{AB(2)} + \left(\boldsymbol{Z}_{i,AB}\right)^T \boldsymbol{\gamma}_i$$

$$\mu_{i,AB(3)} = \alpha_{AB(3)} + \left(\boldsymbol{X}_{i,AB} - \boldsymbol{X}_{Ref,AB}\right)^T \boldsymbol{\beta}_{AB(3)} + \left(\boldsymbol{Z}_{i,AB}\right)^T \boldsymbol{\gamma}_i$$

where

$\alpha_{AB(1)}, \alpha_{AB(2)}, \alpha_{AB(3)}$ are intercept parameters
$\boldsymbol{X}_{i,AB}$ is the $r \times 1$ fixed-effects design vector
$\boldsymbol{X}_{Ref,AB}$ is the $r \times 1$ fixed-effects reference design vector
$\boldsymbol{\beta}_{AB(1)}, \boldsymbol{\beta}_{AB(2)}, \boldsymbol{\beta}_{AB(3)}$ are $r \times 1$ parameter vectors
$\boldsymbol{Z}_{i,AB}$ is the $s \times 1$ random-effects design vector
$\boldsymbol{\gamma}_i$ is the $s \times 1$ random-effects parameter vector

The AsthmaNet BARD Trial

- $\boldsymbol{X}_{i,AB}$, the $r \times 1$ fixed-effects design vector for the A vs B comparison within the i^{th} BARD participant, $i = 1,2, \ldots, I$, can contain effects for period, sequence, clinical center, demographics, lifestyles, genetics, baseline biomarkers, etc.

- Assume that $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_I \; i.i.d. \sim N_s(\boldsymbol{0}, \boldsymbol{\Gamma})$. A reasonable approach is to set

$$\boldsymbol{\gamma}_i = \begin{bmatrix} \gamma_{i,A} \\ \gamma_{i,B} \\ \gamma_{i,C} \\ \gamma_{i,D} \end{bmatrix}, \boldsymbol{\Gamma} = \begin{bmatrix} \sigma_{AA} & 0 & 0 & 0 \\ 0 & \sigma_{BB} & 0 & 0 \\ 0 & 0 & \sigma_{CC} & 0 \\ 0 & 0 & 0 & \sigma_{DD} \end{bmatrix}, \text{ and}$$

$$\left(\boldsymbol{Z}_{i,AB}\right)^T = [1\ 1\ 0\ 0], \left(\boldsymbol{Z}_{i,AC}\right)^T = [1\ 0\ 1\ 0], \text{ etc.}$$

The AsthmaNet BARD Trial

- Obviously, the regression model for the first logit is of primary importance because it focuses directly on the superiority of one treatment regimen over another within a pairwise comparison.

- The quadrinomial logistic regression model with random effects, however, also is appealing because it provides numerous opportunities for exploratory data analyses of those regressors that might impact
  - treatment superiority (first logit)
  - treatment preference (second logit)
  - missingness (third logit)

The AsthmaNet BARD Trial

- The population-level odds ratios for the pairwise treatment comparisons at the reference levels are

$$exp\left(\alpha_{AB(1)}\right),\ exp\left(\alpha_{AC(1)}\right),\ exp\left(\alpha_{AD(1)}\right),\ exp\left(\alpha_{BC(1)}\right),$$
$$exp\left(\alpha_{BD(1)}\right),\ exp\left(\alpha_{CD(1)}\right)$$

- If the researchers desire estimated probabilities for individual participants, then

$$\Pr\left[Y_{i,AB} = +1\right] = \mu_{i,AB(1)}\mu_{i,AB(2)}\mu_{i,AB(3)}$$

$$\Pr\left[Y_{i,AB} = -1\right] = \left\{1 - \mu_{i,AB(1)}\right\}\mu_{i,AB(2)}\mu_{i,AB(3)}$$

$$\Pr\left[Y_{i,AB} = 0\right] = \left\{1 - \mu_{i,AB(2)}\right\}\mu_{i,AB(3)}$$

$$\Pr\left[Y_{i,AB} = .\right] = \left\{1 - \mu_{i,AB(3)}\right\}$$

The AsthmaNet BARD Trial

- Thus, not only does the BARD trial attempt to identify the optimal treatment regimen for each individual study participant, it also provides the framework for assessing population-level comparisons of the treatment regimens.

- Other large clinical trials could mimic the BARD design, so it is possible that a series of N-of-1 trials might have a role in Phase III clinical research.

Summary

- An N-of-1 clinical trial is valuable for identifying optimal treatments within an individual patient. It is an important tool in precision medicine.

- A typical N-of-1 clinical trial invokes a multi-way crossover design, and a one-lag autoregressive model is a reasonable approach for statistical analysis.

- Meta-analytic methods can be applied to combine the results from a series of N-of-1 trials. Bayesian models are especially useful and prior information should be shared across individuals.

- A CONSORT statement is available for reporting the design, conduct, and analysis of N-of-1 trials.

Summary

- The AsthmaNet “Best African American Response to Asthma Drugs (BARD)” trial invoked a four-way crossover design and can be considered a series of N-of-1 trials.

- A multinomial logistic regression model with random effects is in progress to combine results across all BARD participants to investigate population-level treatment comparisons.

- Multi-way crossover trials, in the manner of the BARD trial, should be considered if the clinical researchers have dual objectives of
  - investigating optimal treatment regimens for individual patients
  - assessing population-level treatment effects

REFERENCES

- Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Personalized Medicine* 2011; 8:161-173.
- Mahon JL, Laupacis A, Hodder RV, McKim DA, Paterson NAM, Wood TE, Donner A. Theophylline for irreversible chronic airflow limitation: A randomized study comparing n of 1 trials to standard practice. *Chest* 1999; 115:38-48.
- 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. *Journal of Clinical Epidemiology* 1997; 50:401-410.
- 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; 63:1312-1323.
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis, Third Edition*. Boca Raton, FL: CRC Press, 2013.
- Schluter PJ, Ware RS. Single patient (n-of-1) trials with binary treatment preference. *Statistics in Medicine* 2005; 24:2625-2636.
- Oleson JJ. Bayesian credible intervals for binomial proportions in a single patient trial. *Statistical Methods in Medical Research* 2010; 19:559-574.
- Vohra S, Shamseer L, Sampson M, Bukutu C, Schmid CH, Tate R, Nikles J, Zucker DR, Kravitz R, Guyatt G, Altman DG, Moher D, and the CENT group. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ* 2015; 350:h1738 doi: 10.1136/bmj.h1738.