Chapter

Optimal N-of-1 Clinical Trials for Individualized Patient Care and Aggregated N-of-1 Designs

Yin Li, Weng Kee Wong and Keumhee Chough Carriere

Abstract

Precision medicine typically refers to the use of genomic signatures of patients to assign more effective therapies to treat patients, or, for improved diagnosis of the early onset of a disease so that interventions can be delivered to prevent or delay the disease progression. Because the aim is to provide individualized patient treatment, such single-person trials are called N-of-1 trials. This chapter reviews fundamental ideas, models, and construction of optimal designs for N-of-1 trials, which are invariably constructed from crossover trials, where each patient receives a random sequence of trial treatments over time. We construct examples of universally optimal N-of-1 designs for comparing two treatments under various correlation structure assumptions and discuss how N-of-1 trials may be combined to form optimal aggregated N-of-1 trials for assessing average treatment effects for two or more treatments.

Keywords: crossover design, individualized care, N-of-1 trials, precision medicine, universally optimal designs

1. Introduction

N-of-1 trials or single-patient trials focus on one patient and their main goal is to evaluate whether the treatment is effective for the individual. The main motivation for such trials is that each patient serves as his or her own control, and another is that each patient is different from another and there is no average patient. This is in contrast to conventional clinical trials where the aim is to optimize treatment for the average patients. Consequently, their aims are different, and conventional clinical trials are not appropriate for N-of-1 trials. These trials may appear new but they are not, except that they probably were given short shrift and not well publicized. In the last decade or so, there is increasing interest in N-of-1 trials. Duan et al. [1] raised awareness among clinicians and epidemiologists that N-of-1 trials are potentially useful for informing personalized treatment decisions for patients with chronic conditions. A monograph on this topic in healthcare is [2], where their applications to behavioral sciences and many medical settings are discussed, including the economics, ethics, statistical analysis of running such trials, and how to report results to professional audiences. Scuffham et al. [3] showed how N-of-1 trials can improve patient

management and save costs and Kravitz and Duan [4] provided a user's handbook on implementing such trials. A systematic review of the use of N-of-1 trials in the medical literature is given in [5]. There are many ways to analyze and compare results from N-of-1 trials; see for example, [6].

Interestingly, and perhaps, not unexpectedly, results from N-of-1 trials can be combined to generate group mean effects, as [7, 8] demonstrated how it can be done using systematic reviews and meta-analyses on the effects of amphetamine and methylphenidate for attention-deficit hyperactivity disorder. Li et al. [9] provided a systematic review of quality N-of-1 trials published between 1985 and 2013 in the medical literature based on the CONSORT extension for N-of-1 Trials (CENT) where they examined factors that influence reporting quality in these trials. In palliative care, Senior et al. [10] designed a N-of-1 trial of a psychostimulant, methylphenidate hydrochloride (MPH) (5 mg bd), compared to placebo as a treatment for fatigue, with a population estimate of the benefit by the aggregation of multiple SPTs. Forty patients who had advanced cancer was enrolled through specialist palliative care services in Australia.

Multi-crossover single-patient trials are often employed when the focus is to make the best possible treatment decision for an individual patient [2, 11, 12]. From a clinician's perspective, having clear evidence of the value of one treatment over another (or no treatment) is more useful than knowing the average response. The average response gives the clinician the probability that a treatment will be effective, whereas N-of-1 trials give more certainty about whether the treatment for a particular patient will work or not.

In what is to follow, we assume that there are predetermined p periods in the crossover study, and in each period only one of the treatments is administrated. The same treatment may be used in other periods. We first discuss the case when there are two treatments and two periods for N-of-1 trials before extending them to aggregated N-of-1 trials to evaluate the effects of treatments for the average patients. Treatment groups are generically denoted by A, B, C, and so on.

Many researchers studied the optimality of crossover designs [13–18]. Optimal designs have been constructed under a variety of statistical models to provide the most accurate inference of the treatment effects. It is known that the two-treatment design AB, AA and their duals BA, BB are found to be universally optimal for two-period experiments, with the duality defined as the sequence that switches A and B with the same effect. Similarly, it is known that the two-sequence design ABB and its dual BAA and the four-sequence design $ABBA$, $AABB$ and their duals $BAAB$, $BBAA$ are optimal for three- and four-period experiments, respectively [19] and [20].

A direct application of this two-treatment optimal design results from the literature with A replaced as AB and B as BA would suggest that optimal N-of-1 trials can use the four-sequence design with $ABBA$, $ABAB$ or their duals for two within-patient comparisons. Similarly, the two-sequence design with $ABBABA$ or its dual may be optimal for three within-patient comparisons, and the four-sequence design with $ABBABAAB$, $ABABBABA$ or any of their duals may be optimal for four within-patient comparisons.

However, design issues are not always as straightforward to address. For example, Carriere and Li [21] showed that constructing N-of-1 trials for individualized care from sequences in these repeated measurement designs is not always optimal for estimating individual-based treatment effects. Likewise, Guyatt et al. [22] showed that aggregating a series of N-of-1 trials that are optimal for individual patients can also provide an optimal estimate of the treatment effects for the average patient. For example, in a multi-clinic setting in three AB pair six-period N-of-1 studies, all eight possible sequences ($2^{6/2} = 8$) have been used, i.e., $ABABAB$, $ABABBA$, $ABBAAB$, $ABBABA$ and

their duals to estimate both individual-based and average treatment effects. However, we show how these do not lead to optimal aggregated N-of-1 trials for estimating the treatment effects for the average patient.

2. Models and information matrix

The traditional crossover design model assumes that the carryover effects last for only one period. The patient effects are considered fixed in the model. The traditional model assumes no carryover effects for the observations in the first period. An alternative model which has carryover effects in the first period as well is built by giving patients a pre-period or baseline period [23, 24]. More complex models have also been considered. Some models incorporate higher-order carryover effects [25]; some consider carryover effects are proportional to the treatment effects [26], some include the interaction effects between the treatment effects and carryover effects [27], and others have random patient effects [28–30].

We first focus on the traditional model, frequently used to analyze repeated measures crossover data:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \tau_{d(i,j)} + \gamma_{d(i-1,j)} + \varepsilon_{ij}, \tag{1}$$

$i = 1, \cdots, p$ and $j = 1, \cdots, N$. Here Y_{ij} is the outcome in the i^{th} period from the j^{th} patient; α_i is the i^{th} period effect and β_j is the j^{th} patient effect. Further, $d(i, j)$ represents the treatment assigned to the patient in period i of patient j, and $\tau_{d(i,j)}$ and $\gamma_{d(i-1,j)}$ are, respectively, the treatment effect of the treatment in the i^{th} period and the carryover effect of the treatment in the $(i-1)^{th}$ period.

The model assumes that the carryover effects only depend on the treatment assigned in the previous period but not on the treatment in the current period, which may be unrealistic. Taking the interaction into account without introducing too many parameters, Kunert and Stufken [17] presented a model with self and mixed carryover effects. The self carryover effect occurs when the treatments administered in the current and the previous periods are the same; otherwise, we have a mixed carryover effect. The model with the self and mixed carryover effects is given by

$$Y_{ij} = \begin{cases} \mu + \alpha_i + \beta_j + \tau_{d(i,j)} + \gamma_{s,d(i-1,j)} + \varepsilon_{ij}, & \text{if d(i, j) = d(i-1, j)} \\ \mu + \alpha_i + \beta_j + \tau_{d(i,j)} + \gamma_{m,d(i-1,j)} + \varepsilon_{ij}, & \text{if d(i, j) } \neq \text{ d(i-1, j)} \end{cases}, \tag{2}$$

where α_i, β_j, $d(i, j)$ and $\tau_{d(i,j)}$, are defined as in model (1). The parameters $\gamma_{s,d(i-1,j)}$ and $\gamma_{m,d(i-1,j)}$ represent the self and mixed carryover effects of the treatment assigned in the $(i-1)^{th}$ period, respectively.

In an N-of-1 trial with $N = 1$, the j index can be omitted. Further, with one patient and p responses in total, the period effects and patient effects cannot be accommodated. Therefore, we need to reduce the models for the case when $N = 1$.

For models (1) and (2), we define the contrast of the direct treatment effects by $\tau = (\tau_A - \tau_B)/2$, the contrast of the first-order carryover effects by $\gamma = (\gamma_A - \gamma_B)/2$, the contrast of the self carryover effects by $\gamma_s = (\gamma_{s,A} - \gamma_{s,B})/2$ and the contrast of the mixed carryover effects by $\gamma_m = (\gamma_{m,A} - \gamma_{m,B})/2$.

To construct a model-based optimal design, we commonly use design criteria such as A-, D-, and E-optimality. The A-, D-, and E-optimal design maximizes the trace, the determinant, or the eigenvalue of the information matrix among a class of all competing designs. The information matrix measures the amount of information about the unknown model parameters. Formally, given the model and the design, the elements in the information matrix are found by first taking the expectation of the second derivatives of the complete log-likelihood function with respect to the parameters and multiplying them by -1. In practice, not all model parameters are of interest. In this case, we would first partition the information matrix and work with the submatrix corresponding to the parameters of interest.

Specifically, we first partition the design matrix $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$, where $\mathbf{X}_1$ contains the columns of the design matrix pertaining to nuisance parameters and $\mathbf{X}_2$ contains columns corresponding to the parameters of interest. The vector of model parameters θ is likewise partitioned as $\theta = (\tau,\gamma)'$ or $(\tau, \gamma_s, \gamma_m)'$, representing the direct treatment effects and carryover effects. Then, with Σ denoting the covariance matrix, the information matrix can be written as

$$I_d(\theta) = X_2'\Sigma^{-1}X_2 - X_2'\Sigma^{-1}X_1\left(X_1'\Sigma^{-1}X_1\right)^{-1}X_1'\Sigma^{-1}X_2. \tag{3}$$

Following [13], a design is universally optimal if (i) its information matrix is completely symmetric, and (ii) it maximizes the trace of the information matrix. To study the universal optimality of treatment effects in the t treatments N-of-1 designs, we obtain the information matrix for the parameters of interest under the traditional model. Then the universally optimal designs could be constructed as long as the conditions given by [13] are satisfied.

2.1 Cycles and sequences

We first discuss how to find N-of-1 designs for comparing two treatments by minimizing the variance of the estimated direct treatment effect contrast, τ. To this end, it is helpful to define sequence feature parameters and show the association between them and the sequences in N-of-1 designs is useful for finding optimal N-of-1 trials for model (1) and (2) .

For N-of-1 trials involving two treatments, the design sequences consist of crossover pairs, AB and BA. Within each crossover pair, the two treatments are distinct. For two consecutive crossover pairs, the treatments assigned to the second period in the previous pair and the first period in the latter pair can be different or the same.

Further, if an AB pair is followed by a BA pair, as in $ABBA$ (or $BAAB$), we define the design as having alternating pairs in the sequence. The performance of an N-of-1 trial sequence is related to how the pairs AB and BA alternate. The following feature parameters define how AB and BA alternate in a sequence.

- s: the number of subsequences of AA and BB;
- m: the number of subsequences of AB and BA;
- $h = s - m$: the indicator of how often treatments crossed between subsequences .

When we define s and m, the subsequences can be constructed by either the treatments from a crossover pair, or be the treatments assigned to the second period in

h	**s**	**m**
$-(p-1)$	0	p-1
$-(p-3)$	1	p-2
$-(p-5)$	2	p-3
⋮	⋮	⋮
-1	$\frac{p}{2}-1$	$\frac{p}{2}$

Table 1.
Feature parameters of a sequence in a 2-treatment N-of-1 design.

h	**Sequence**	**Alternation**	***s***	***m***
−7	*ABABABAB*	0	0	7
−5	*ABABABBA*	1	1	6
	ABABBABA		1	6
	ABBABABA		1	6
−3	*ABABBAAB*	2	2	5
	ABBAABAB		2	5
	ABBABAAB		2	5
−1	*ABBAABBA*	3	3	4

Note: s = the number of AA and BB;m = the number of AB and BA in a treatment sequence, and h = s − m.

Table 2.
Sequences for $p = 8$ with corresponding design parameter values.

the previous pair and the first period in the latter pair. Therefore, in a p-period sequence, there are $p-1$ such subsequences with a length of 2. By the definition of feature parameters, we have $s+m=p-1$. Determined by how a sequence is constructed, the value of h is negative and takes on possible values in $-1, -3, \cdots, -(p-1)$. **Table 1** displays the relationship among h, s and m.

For a particular h, we calculate s and m by setting $s=(p-1+h)/2$ and $m=(p-1-h)/2$. Further, for any given p, the N-of-1 designs can be classified by h. As an example, for $p=8$, **Table 2** shows the relationship between the design sequences and the feature parameters.

In the next section, we show that the information matrix of the parameters of interest are only dependent on the feature parameters. That is, sequences with the same h values have the same information matrix. For instance, when $h=-3$, the three sequences *ABABBAAB*, *ABBAABAB*, *ABBABAAB* and their dual sequences share the same information matrix. If this h is the optimum value, the 8 period N-of-1 trials can use any of these three sequences and their duals.

3. Optimal 2-treatment N-of-1 designs

Let $\mathbf{x}_\tau$, $\mathbf{x}_\gamma$, $\mathbf{x}_s$ and $\mathbf{x}_m$ be the design vectors corresponding to the parameters τ, γ, γ_s and γ_m, respectively. Under the traditional model, the design matrix is $[\mathbf{1}_p, \mathbf{x}_\tau, \mathbf{x}_\gamma]$ for the parameters $[\mu, \tau, \gamma]$ with $\mathbf{X}_1 = 1_p$ and $\mathbf{X}_2=[\mathbf{x}_\tau, \mathbf{x}_\gamma]$. Under the self and mixed effect

model, we have $[\mathbf{1}_p\ \mathbf{x}_\tau, \mathbf{x}_s, \mathbf{x}_m]$ for the parameters $[\mu, \tau, \gamma_s, \gamma_m]$ with $\mathbf{X}_1 = 1_p$ and $\mathbf{X}_2 = [\mathbf{x}_\tau, \mathbf{x}_s, \mathbf{x}_m]$.

In 2-treatment N-of-1 trials, the $I_d(\tau, \gamma)$ is a function of the quantities:

$$x'_\tau x_\tau = p, \quad x'_\gamma x_\gamma = p - 1, \quad x'_\tau x_\gamma = h \tag{4}$$

under model (1) for $\theta = (\tau, \gamma)'$, or

$$x'_\tau x_s = s, \quad x'_s x_s = s, \quad x'_m x_s = 0, \quad x'_\tau x_m = -m, \quad x'_m x_m = m \tag{5}$$

under model (2) for $\theta = (\tau, \gamma_s, \gamma_m)'$. Hence, the information matrix can be expressed in terms of p, s, m and h.

Since the information matrices can be further simply expressed in terms of h and p only, for a given p, the optimal $p-$period N-of-1 trial is completely determined by h, and much simpler to construct than previously. We proceed by defining $I_d(\tau)$ appropriately to find designs that maximize the information below.

Under an equi-correlated error assumption, the optimal N-of-1 trial for τ and γ is the one sequence design that consists of pairs of AB and BA appearing alternatively. Hence, the optimal N-of-1 trials for 4, 6, and 8 periods are the one sequence design, $ABBA$, $ABBAAB$ and $ABBAABBA$, respectively. One could switch A and B to obtain a dual sequence with the same effect.

Under the equi-correlated errors, the optimal N-of-1 trial for estimating the direct treatment contrast is the sequence with only AB (BA) pairs with no alternation, such as $BABABA$ and $ABABABAB$. A closed form for the optimal h is complicated for autoregressive errors, and selected numerical results are found when $h = 1 - p$.

To summarize, the optimal N-of-1 trials for estimating direct treatment effects are determined by the three feature parameters h, s, and m. However, specifying one of these along with p determines the design sequence, as illustrated in **Table 2**. We used h to summarize the optimal designs under both the traditional and self and mixed models for 4, 6, 8, 10, and 12-period N-of-1 trials.

It can be shown that under the traditional model, the optimal trial for the direct treatment effect uses the sequence with $h = -1$ for all covariance structures. Therefore, the optimal N-of-1 trial for estimating the direct treatment effect is to alternate between AB and BA pairs. In case that the carryover effect is of interest, it can be easily shown that these designs are also optimal for estimating the carryover effect, which can be obtained using the same technique for optimal designs in treatment effects. Under the self and mixed effects model, the optimal N-of-1 trial for the direct treatment effect uses a sequence with $h = -(p - 1)$ for both uncorrelated and equal-correlated covariances. Therefore, the optimal N-of-1 trial is to use only AB pairs throughout. Under the auto-regressive covariance structure, however, the optimal designs depend on the value of p and the auto-regressive correlation ρ. Generally, the optimal design uses AB and BA pairs alternately, but as ρ or p increases, some abnormalities are observed.

4. Optimal aggregated N-of-1 trial designs with $N > 1$

In addition to the interest in the patient-based evidence of a treatment contrast, it may also be desirable to obtain a population average effect of treatments.

Aggregating the series of N-of-1 trials can give such an estimate of the average effect [7]. Using the one sequence that was found optimal for N-of-1 trial to all patients seems to be an obvious choice. However, it might not optimize the trial for estimating the effects on the average patient and therefore, using the one sequence that is optimal for a single individual patient to all patients might not serve this purpose.

The optimal designs for aggregated N-of-1 trials can also be derived from the information matrices we obtained, similarly as for N-of-1 trials for one patient, by allowing $j = 1, \ldots, N$ with $N > 1$. We approached the problem in two steps; first, we optimize single N-of-1 trials, as the primary goal is to optimize estimating the effects for each patient. Next, we optimize the overall N-of-1 trials in aggregation.

To find the optimal design, we typically choose N_k for $k = 1, \ldots, s$ to allocate patients to a sequence s. The sufficient condition on N_k was given by [20] for a design to be optimal. The condition is called a duality in the design matrices, as defined earlier. Among other things, it permits simplification of the search for the optimal choice for N_k (see also [16]).

As noted earlier for **Table 1**, designs with the same value of h perform equally in estimation precision. Although all or only one of those with an equally optimal h can be used in a trial, practical consideration will lead to using the least necessary number of sequences for ease of treatment administration. Further, we found that there is a unique N-of-1 trial sequence in all p-period experiments. Since the designation of A and B is arbitrary, the optimal N-of-1 trial can be obtained by reversing the order of treatment administration. For example, the optimal 6-period N-of-1 trial is *ABBAAB* under the traditional model for $N = 1$. Its dual, *BAABBA* also has the same value of $h = -1$ and is optimal. Hence, when $N = 1$, either of these two sequences will provide the maximum amount of information. When $N > 1$ and a multiple of 2, we can adopt both of these sequences, as they maximize the information, and this approach also simplifies the search for the optimal design for estimating the treatment effect for the average patients, satisfying the duality condition in [20]. Based on this rationale, we make the following two propositions.

Proposition 1: The optimal design for aggregated N-of-1 trials under the traditional model is to allocate the same number of patients to the optimal sequence with *AB* and *BA* alternating and its dual.

For example, the optimal design for aggregated six-period N-of-1 trials is the two-sequence design using sequences *ABBAAB* and *BAABBA*, allocating the same number of patients to each. For a balanced design, N must be a multiple of 2.

Proposition 2: The optimal design for aggregated N-of-1 trials under the self and mixed model is to allocate the same number of patients to the optimal sequence with no alternation between *AB* and *BA* pairs and its dual. However, under the autocorrelation errors, the optimal design is to allocate the same number of patients to the optimal sequence that alternates between *AB* and *BA* pairs subsequently and its dual.

For example, the optimal aggregated 6-period N-of-1 trials for multi-clinic setting is to use the two-sequence design *ABABAB* and *BABABA* under the equal or uncorrelated errors, and to use the two-sequence design *ABBAAB* and *BAABBA* under the autocorrelated errors, allocating the same number of patients to each sequence.

From each sequence, we obtain individual patient specific treatment effects and by aggregating these one sequence of N-of-1 trials, we can quantify the average treatment effects.

4.1 Numerical comparisons

To appreciate the practical performance of the optimal N-of-1 trials we constructed, we compare the efficiencies of selected designs for estimating the treatment and carryover effects under the two models. We also investigate their performances in some aggregation for estimating the average treatment effect. We limit the comparison to the models with independent and equi-correlation errors. In our comparisons, we also reference many designs, labeled with an A or S at the beginning, like A65 and S83, that were investigated in [31].

Recall that the optimal N-of-1 trials are either to alternate between *AB* and *BA* pairs or simply to repeat the *AB* pair in a sequence. Under the traditional model, the optimal N-of-1 trial uses *ABBAAB* and *ABBAABBA* for 6 and 8 period experiments, respectively. We refer them to S63 and S83. Under the self and mixed effects model, the optimal N-of-1 trial is to use *ABABAB* and *ABABABAB* for 6 and 8 period experiments, respectively, which we refer to S61 and S81. **Table 3** considers other mixtures and shows that the optimal individual-based N-of-1 trials are S63 and S81 under the respective models, as expected. We also observe from the table that (i) there are no real practical differences among various N-of-1 trials under the self and mixed model, and (ii) designs S61 and S81 cannot estimate self carryover effects, making S63 and S83 preferable. Therefore, we recommend using a sequence that alternates between *AB* and *BA* pairs, such as S63 and S83, as robust and optimal N-of-1 trials for all models.

Based on these single sequence trials, we also consider aggregated N-of-1 trials to numerically justify Propositions 1 and 2. We constructed 5 aggregated N-of-1 trials for $p = 6$ and $p = 8$ with $N = 32$ and compare their efficiencies for estimating the average treatment effects as follows.

- A61. *ABABAB* and its dual with 16 patients in each sequence
- A62. *ABABBA* and its dual with 16 patients in each sequence
- A63. *ABBAAB* and its dual with 16 patients in each sequence
- A64. *ABBAAB*, *ABABBA* and their duals with 8 patients in each sequence
- A65. All 8 sequences, S61–S64 and their duals with 4 patients in each sequence
- A81. *ABABABAB* and its dual with 16 patients in each sequence
- A82. *ABABBABA* and its dual with 16 patients in each sequence
- A83. *ABBAABBA* and its dual with 16 patients in each sequence
- A84. *ABABBABA*, *ABBABAAB* and their duals with 8 patients in each sequence
- A85. All 8 sequences, S81–S84 and their duals with 4 patients in each sequence

The design A63 uses the optimal sequence S63 under the traditional model; the design A61 uses the optimal sequence S61 under the self and mixed model although the self carryover effect is not estimable; the design A62 is a slight rearrangement of

Design	h	Traditional model		Self and mixed model		
		$\text{var}(\hat{\tau})$	$\text{var}(\hat{\gamma})$	$\text{var}(\hat{\tau})$	$\text{var}(\hat{\gamma}_s)$	$\text{var}(\hat{\gamma}_m)$
S61: ABABAB	−5	1.208	1.500	1.208	NE	1.500
S62: ABABBA	−3	0.242	0.300	1.250	3.000	1.500
S63: ABBAAB	−1	0.173	0.214	1.214	1.714	1.714
S64: ABBABA	−3	0.242	0.300	1.250	3.000	1.500
A61 = S61 + dual		1.208	1.500	1.208	NE	1.500
A62 = S62 + dual		0.242	0.300	1.250	3.000	1.500
A63 = S63 + dual		0.173	0.214	1.214	1.714	1.714
A64 = S63 + S62 + duals		0.193	0.240	1.210	2.063	1.563
A65 = S61:S64 + duals		0.242	0.300	1.214	2.535	1.521
S81: ABABABAB	−7	1.146	1.333	1.146	NE	1.333
S82: ABABBABA	−5	0.229	0.267	1.167	2.667	1.333
S83: ABBAABBA	−1	0.127	0.148	1.150	1.600	1.400
S84: ABBABAAB	−3	0.150	0.174	1.147	1.647	1.412
A81 = S81 + dual		1.146	1.333	1.146	NE	1.333
A82 = S82 + dual		0.229	0.267	1.167	2.667	1.333
A83 = S83 + dual		0.127	0.148	1.150	1.600	1.400
A84 = S82 + S84 + dual		0.176	0.205	1.147	1.945	1.358
A85 = S81:84 + dual		0.176	0.205	1.147	1.945	1.358

Note: NE means "Not Estimable." For $N = 1$, a six-period N-of-1 trial may consider any one of S61,⋯, S64. For $N > 1$, aggregated six-period N-of-1 trials may use a combination of these, A61,⋯, A65. Similarly, an eight-period N-of-1 trial may consider any one of S81,⋯, S84. For $N > 1$, aggregated six-period N-of-1 trials may use a combination of these, A81,⋯, A85. The variances reported are divided by σ_ε^2/N (under an independence error) or $\sigma_\varepsilon^2(1-\rho)/N$ (under an equi-correlated error).

Table 3.
Variances of the estimators of treatment and carryover effects in six- and eight-period designs.

designs A61 and A63; the design A64 is a combination of designs A62 and A63; the design A65 contains all 8 possible sequences of a 6-period design. Designs A81–A85 are also similarly constructed from various N-of-1 trials. We compare these designs under the traditional model and the self and mixed model. **Table 3** displays the comparison results under the two models and reports the variances of the estimated τ, γ, γ_s and γ_m after they are divided by their leading constants σ^2/N (when errors are independent) or by $\sigma^2(1-\rho)/N$ (when errors equi-correlated).

Table 3 shows that under the traditional model, the design A63 with the optimal sequence *ABBAAB* and its dual provides the best precision for estimating both the direct treatment effect and the carryover effect for the average patients. Each of the sequences optimally estimates the individual-based treatment effect. The least efficient choice would be the design A61. Design A65, which has been used in a recent multi-clinical trial [7], is rather inefficient as well, not to mention the unnecessarily lengthy administration time and cost required to manage many treatment groups, which requires the number of patients to be a multiple of 8.

When using the self and mixed effects model, Design A61 provides the best precision for estimating the direct treatment effect and the mixed carryover effect. However, the self carryover effect is not estimable. Overall, A63 is the optimal choice even in this case. However, all designs performed rather similar with over 95% relative efficiency under the self and mixed effects model, as observed earlier for single N-of-1 trials.

A similar observation is possible for 8-period designs and their sequences. In summary, it appears that there is no discernable advantage to distinguish among the two models and various error structures.

Overall, S63 and S83 for single N-of-1 trials or designs A63 and A83 in aggregation of S63, S83 and their duals seem to be the best under both models. They are optimal for estimating direct treatment and mixed carryover effects. Further, they are optimal for estimating both the treatment and carryover effects under the traditional model. Hence, we conclude that the optimal six-period aggregated N-of-1 trials is *ABBAAB* and its dual *BAABBA*, while the optimal eight-period aggregated N-of-1 trials is *ABBAABBA* and its dual *BAABBAAB*. For an N-of-1 trial, using one of these sequences will optimize the treatment for an individual patient.

We close this section with a summary note. Our numerical work suggests that alternating *AB* and *BA* pairs in sequence is likely to result in an optimal or nearly optimal $p-$period design for all the models we have considered for estimating both individual effects in N-of-1 trials and average effects in aggregated N-of-1 trials.

5. Universally optimal N-of-1 designs for more than two treatments

Oftentimes, N-of-1 trials deal with comparing $t > 2$ treatments and we briefly discuss selected universal optimal N-of-1 trials for such a situation. In N-of-1 trials with $t > 2$ treatments, we can consider a sequence consisting of treatments in blocks of a size t. Every block within the sequence contains each of the t treatments exactly once. It follows that N-of-1 designs constructed in this way can ensure treatments are compared fairly, and poor balance can be prevented when the study is terminated prematurely. For example, in a 3-treatment N-of-1 trial, a six-period design could be *ABC*|*BCA*, where the sign | divides them into blocks. Li [31] denoted such a class of N-of-1 trial designs by No1(t,t), where the first t in the notation represents the number of treatments in the study and the second t denotes that the treatments are to be administered to be in blocks of size t. Therefore, a six-period design in the above example is No1(3,3).

Li [31] showed that Kiefer's conditions could not be satisfied with designs in the class No1(t,t). However, if we consider a slightly different class of designs, then universally optimal designs can be obtained. Li [31] used No1(t,t+) to denote the new class of designs, which consist of designs with one extra treatment to the last period in No1(t,t). For instance, a design in No1(3,3+) could be *ABC*|*BCA*|*A*, *ABC*|*BCA*|*B*, or *ABC*|*BCA*|*C*. Similarly, some examples from the class No1(2,2+) are *AB*|*BA*|*A*, *AB*|*BA*|*B*, etc.

One disadvantage of the universally optimal designs for $t > 2$ treatments is that the length of the sequence can be unmanageable, leading to drop-outs and non-compliances before the end of the trials. As discussed in [3], a universally optimal design in No1(t,t+) requires the length of the sequence equal to $(t-1)!t^2+1$, which is 5 for $t = 2$, 19 for $t = 3$, and 97 for $t = 4$. It may be infeasible in practice because the longer the period of the experiment, the more expensive the experiment and the

higher the risk of drop-outs. To shorten the length of the experiment without losing the balance in the comparison of treatments, Li [31] introduced a class of designs in No1(t,s) or No1(t,s+) for some $s < t$, especially when $s = 2$. To do so, the restriction that the block size must be equal to the number of the treatments can be relaxed [31]. By allowing the block size to be smaller than t, universally optimal designs can be manageable in practice, thereby reducing the risk of early dropouts and the burden of treatment administration.

Li [31] showed some practical universally optimal designs for three-, four- and five-treatment in blocks of size 2. In each block, two different treatments are assigned such as a crossover pair. For t-treatment designs, there are $t(t-1)$ different kinds of crossover pairs. To construct the universally optimal design, the crossover pairs are selected such that each subsequence of A_iA_j, $1 \le i, j \le t$, appears only once. Therefore, for universally optimal designs, the number of periods is $p = t^2 + 1$. For example, p is 10 for three-treatment designs, 17 for four-treatment designs and 26 for five-treatment designs. We close by giving examples of universally optimal designs in selected situations. Omitting details, which are available in [31], they are:

- No1(3,2) with $t = 3$:

$$\{ABBCCAACBA\} \quad or \quad \{BCABBAACCB\},$$

- No1(4,2+) for $t = 4$:

$$\{ABBCCDDACBDCADBAA\} \quad or \quad \{BCCDDAABBDCADBACB\}$$

and

- No1(5,2) for $t = 5$:

$$\{ABBCCDDEEAACBDCEDAEBADBECA\} \quad or$$
$$\{CDDEEAABBCCEBDACBEDBAECADC\}.$$

6. Concluding remarks

In this Chapter, we discussed and reviewed construction of universally optimal N-of-1 designs and how they may be aggregated to estimate treatment effects for the average patients. Originally, Kiefer [13] proposed the concept of universal optimality with zero row and column sums in the information matrices. We examined conditions when such universally optimal designs exist with special application to N-of-1 trial designs that will make them optimal no matter what criteria are applied. In particular, we first presented a sufficient condition that ensures N-of-1 designs are universally optimal for the traditional model that accommodates the carryover effects. Additionally, we discussed extensions of our work to finding optimal aggregated N-of-1 designs. Using numerical results from our simulation for comparing the estimated precision of several six- and eight-period designs, we were able to obtain realistic guidelines for the practitioners.

Overall, there are three key conclusions from this chapter. The first is that alternating between *AB* and *BA* pairs in sequence will result in an optimal or nearly optimal N-of-1 trial for a single patient for models considered in this chapter. In particular, our work suggests that alternating between *AB* and *BA* pairs in a single trial is quite robust to mis-specification in the error structures considered in the chapter. Consequently, there is less need to guess or conduct a pilot study to verify model assumptions and the error structures.

Another take home message is that when an experiment has been carried out with the optimal N-of-1 trial and additional patients are accrued in the trial, we can aggregate these N-of-1 trials optimally by allocating the same number of patients to its dual sequence, thereby optimizing the trial for both the individual and average patients.

Lastly, we also provided a strategy for finding N-of-1 trials with more than 2 treatments. By restricting the class of designs and utilizing each subsequence, we constructed universally optimal N-of-1 trial designs when there are $t = 3$, 4, or 5 treatments.

Author details

Yin Li[1], Weng Kee Wong[2] and Keumhee Chough Carriere[3*]

1 Ontario Medical Association, Toronto, ON, Canada

2 Department of Biostatistics, University of California-Los Angeles, LA, CA, USA

3 Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada

*Address all correspondence to: kccarrie@ualberta.ca

References

[1] Duan N, Kravitz R, Schmid C. Single-patient (n-of-1) trials: A pragmatic clinical decision methodology for patient-centered comparative effectiveness research. Journal of Clinical Epidemiology. 2013;**66**:S21-S28

[2] Nikles CJ, Clavarino AM, Del Mar CB. Using N-of-1 trials as a clinical tool to improve prescribing. British Journal of General Practice. 2005;**55**:175-180

[3] Scuffham PA, Nikles J, Mitchell GK, Yelland MJ, Vine N, Poulos CJ, et al. Using N-of-1 trials to improve patient management and save costs. Journal of General Internal Medicine. 2010;**25**(9): 906-913

[4] Kravitz RL, Duan N. Design and Implementation of N-of-1 Trials: A user's Guide. USA: Agency for Healthcare Research and Quality; 2014

[5] Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature a systematic review. Medical Care. 2011;**49**:761-768

[6] Chen X, Chen P. A comparison of four methods for the analysis of N-of-1 trials. PLoS One. 2014;**9**(2): e87752

[7] 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

[8] Punja S, Xu D, Schmidc C, Hartling L. N-of-1 trials can be aggregated to generate group mean treatment effects: A systematic review and meta-analysis. Journal of Clinical Epidemiology. 2016; **76**:65-75

[9] Li J, Gao W, Punja S, Ma B, Vohra S, Duan N, et al. Reporting quality of N-of-1 trials published between 1985 and 2013: A systematic review. Journal of Clinical Epidemiology. 2016;**76**:57-64

[10] Senior HE, Mitchell GK, Nikles J, Carmont SA, Schluter PJ, Currow DC, et al. Using aggregated single patient (N-of-1) trials to determine the effectiveness of psychostimulants to reduce fatigue in advanced cancer patients: A rationale and protocol. BMC Palliative Care. 2013;**12**(1):1-6

[11] Larson EB. N-of-1 clinical trials: A technique for improving medical therapeutics. Western Journal of Medicine. 1990;**152**:52

[12] Edgington ES. Statistics and single case analysis. Progress in Behavior Modification. 1984;**16**:83-119

[13] Kiefer J. General equivalence theory for optimum designs (approximate theory). The Annals of Statistics. 1975;**2**: 849-879

[14] Cheng CS, Wu CF. Balanced repeated measurements designs. The Annals of Statistics. 1980;**8**:1272-1283

[15] Kershner RP, Federer WT. Two-treatment crossover designs for estimating a variety of effects. Journal of the American Statistical Association. 1981;**76**:612-619

[16] Carriere KC, Reinsel GC. Investigation of dual-balanced crossover designs for two treatments. Biometrics. 1992;**48**:1157-1164

[17] Kunert J, Stufken J. Optimal crossover designs in a model with self and mixed carryover effects. Journal of

the American Statistical Association. 2002;**97**:898-906

[18] Kunert J, Stufken J. Optimal crossover designs for two treatments in the presence of mixed and self-carryover effects. Journal of the American Statistical Association. 2008;**103**: 1641-1647

[19] Carriere KC. Crossover designs for clinical trials. Statistics in Medicine. 1994;**13**:1063-1069

[20] Laska EM, Meisner M. A variational approach to optimal two-treatment crossover designs: Application to carryover-effect models. Journal of the American Statistical Association. 1985; **80**:704-710

[21] Carriere KC, Li Y, Mitchell G, Senior H. Methodological considerations for N-of-1 trials. In: The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer. 2015. pp. 67-80

[22] Guyatt GH, Heyting A, Jaeschke R, Keller J, Adachi JD, Roberts RS. N of 1 randomized trials for investigating new drugs. Controlled Clinical Trials. 1990; **11**:88-100

[23] Kunert J. Optimality of balanced uniform repeated measurements designs. The Annals of Statistics. 1984; **12**:1006-1017

[24] Afsarinejad K. Circular balanced uniform repeated measurements designs. Statistics and Probability Letters. 1988;7(3):187-189

[25] Bose M, Mukherjee B. Optimal crossover designs under a general model. Statistics & Probability Letters. 2003;**62**: 413-418

[26] Kempton RA, Ferris SJ, David O. Optimal change-over designs when carry-over effects are proportional to direct effects of treatments. Biometrika. 2001;**88**:391-399

[27] Sen M, Mukerjee R. Optimal repeated measurements designs under interaction. Journal of Statistical Planning and Inference. 1987;**17**:81-91

[28] Saha R. Repeated measurements designs. Calcutta Statistical Association Bulletin. 1983;**32**:153-168

[29] Jones B, Kunert J, Wynn HP. Information matrices for mixed effects models with applications to the optimality of repeated measurements designs. Journal of Statistical Planning and Inference. 1992;**78**:307-316

[30] Carriere KC, Reinsel GC. Optimal two-period repeated measurements designs with two or more treatments. Biometrika. 1993;**80**:924-929

[31] Li Y. Optimal Crossover Designs in Clinical Trials. Canada: University of Alberta; 2017