

RESEARCH ARTICLE

Open Access



# Estimation of treatment preference effects in clinical trials when some participants are indifferent to treatment choice

Stephen D. Walter<sup>1\*</sup>, Robin M. Turner<sup>2</sup>, Petra Macaskill<sup>3</sup>, Kirsten J. McCaffery<sup>3</sup> and Les Irwig<sup>3</sup>

## Abstract

**Background:** In the two-stage randomised trial design, a randomly sampled subset of study participants are permitted to choose their own treatment, while the remaining participants are randomised to treatment in the usual way. Appropriate analysis of the data from both arms of the study allows investigators to estimate the impact on study outcomes of treatment preferences that patients may have, in addition to evaluating the usual direct effect of treatment. In earlier work, we showed how to optimise this design by making a suitable choice of the proportion of participants who should be assigned to the choice arm of the trial. However, we ignored the possibility of some participants being indifferent to the treatments under study. In this paper, we extend our earlier work to consider the analysis of two-stage randomised trials when some participants have no treatment preference, even if they are assigned to the choice arm and allowed to choose.

**Methods:** We compare alternative characterisations of the response profiles of the indifferent or undecided participants, and derive estimates of the treatment and preference effects on study outcomes. We also present corresponding test statistics for these parameters. The methods are illustrated with data from a clinical trial contrasting medical and surgical interventions.

**Results:** Expressions are obtained to estimate and test the impact of treatment choices on study outcomes, as well as the impact of the actual treatment received. Contrasts are defined between patients with stated treatment preferences and those with no preference. Alternative assumptions concerning the outcomes of undecided participants are described, and an approach leading to unbiased estimation and testing is identified.

**Conclusions:** Use of the two-stage design can provide important insights into determinants of study outcomes that are not identifiable with other designs. The design can remain attractive even in the presence of participants with no stated treatment preference.

**Keywords:** Randomised trials, Design, Estimation, Participant preferences

## Background

Clinical trial outcomes may be affected by preferences that trial participants might have between the treatments under comparison. Preference effects can be substantial [1], but they are unobservable in standard trial designs; however, they are estimable in the two-stage randomised trial, in addition to the usual treatment effect [2, 3]. Using this design, a randomly sampled subset of patients

are permitted to choose their own treatment, with the remainder randomised to treatment in the usual way. Estimated preference effects often provide additional insight into determinants of the study response, beyond that of the treatment effect alone [1–4]. As we have reviewed elsewhere [1], several preference-based designs have been employed in biomedical research in recent years [5–18].

We previously showed how to optimise the two-stage design by controlling the proportion of patients who are allocated to the choice arm (and hence can choose their treatment) versus the random arm (where they are

\* Correspondence: walter@mcmaster.ca

<sup>1</sup>Department of Clinical Epidemiology and Biostatistics, McMaster University, CRL 233, Hamilton, ON, Canada L8N 3Z5

Full list of author information is available at the end of the article



randomised to treatment) [3]. In our previous illustrative example, involving the management of women with abnormal cervical screening results [4, 5], the investigators compared a novel intervention with standard care, which effectively defined a treatment preference for all participants. However, in other situations, some participants may have no treatment preference, despite being informed about the alternatives. For instance, in our further example discussed later [6, 7], 69% of participants in the choice arm had no preferred treatment.

In general, patients may be unable or unwilling to choose a treatment because they want the doctor to decide, or through anxiety or a lack of information or decision-making confidence. Note specifically that “no preference” is the *observed* preference value obtained after due process of presenting treatment options to the patient, but it is not “missing” data.

No previous paper has discussed how to deal with patients who are indifferent to treatment choice, i.e., who are “undecided”, in more than passing fashion. Apart from our own previous work [1], parameter estimation in the two-stage design has not been described. The possibility of undecided patients was mentioned when the design was originally proposed by Rucker [2], but only concerning significance testing. Rucker assumed that such participants either did not exist, or that preference effects did not apply to them. We show here that these approaches result in potentially biased parameter estimates, but also that alternative (and testable) assumptions are possible. We believe this issue is an important gap in the existing literature, on a topic with considerable practical applicability.

Our previous work assumed that all patients had a treatment preference, but we now extend our methods to allow for undecided patients. We obtain estimates, variances and test statistics for various preference effects, and estimate the usual treatment effect in this design. We provide a general characterisation of the

undecided group, including alternative assumptions about their response profile, and we determine if these assumptions are empirically testable.

We illustrate the calculations using actual trial data, discuss alternative ways that data from undecided participants can be used, and compare our analysis to that of the original trial investigators. In our Discussion, we consider the appeal of the two-stage design when participants with no treatment preference occur, including their impact on study feasibility and efficiency.

**Methods**

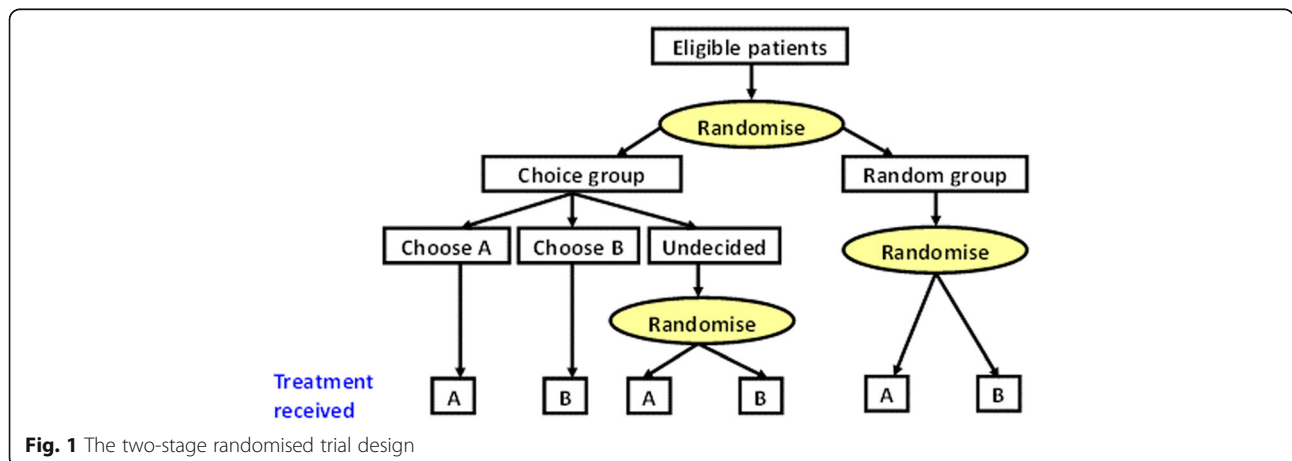
In the two-stage randomised trial design, participants are first randomly divided between the random and choice arms (see Fig. 1). In the random arm, participants are randomised to treatments A or B in the usual way. In the choice arm, participants are allowed to choose their treatment if they have a preference, while participants with no preference are randomised.

Following Rucker [2], we adopt the model:

$$Y_{ijk} = \mu + \tau_i + \nu_j + \pi_{ij} + \varepsilon_{ijk} \tag{1}$$

where  $Y_{ijk}$  is the study outcome for participant  $k$  who receives treatment  $i$  ( $i = 1$  for treatment A, 2 for treatment B) and who is in treatment preference group  $j$  ( $j = 1$  for participants preferring A,  $j = 2$  for participants preferring B, and  $j = 3$  for participants with no preference). The parameters  $\tau$ ,  $\nu$ , and  $\pi$  characterise the treatment, selection and preferences effects, as defined below;  $\varepsilon$  is a random error term, assumed independent of the other terms. As previously, we modify Rucker’s notation by replacing  $\sigma$  with  $\nu$ , to avoid confusion between preference effects and standard deviations that are required later.

To eliminate parameter redundancies, we add four sets of constraints:



$$\tau_1 + \tau_2 = 0,$$

$$\alpha v_1 + \beta v_2 + \gamma v_3 = 0,$$

$$\alpha \pi_{i1} + \beta \pi_{i2} + \gamma \pi_{i3} = 0, \text{ for } i = 1, 2$$

and

$$\pi_{1j} + \pi_{2j} = 0 \text{ for } j = 1, 2, 3.$$

where  $\alpha$ ,  $\beta$  and  $\gamma$  are the expected proportions of participants in the choice arm who choose treatment A, treatment B, or make no choice, respectively.

Table 1 gives our notation. To simplify the expressions and their interpretation for estimates and test statistics to be derived later, we retain  $Y$  as the response variable in the random arm, but use  $X$  for the response in participants with a treatment preference in the choice arm, again following the original formulation [2]. We also denote  $V$  as the outcome among undecided participants in the choice arm. The total sample size is  $N$ , with  $m$  and  $n$  participants being allocated to the choice and random arms. For simplicity we assume equal treatment group sizes in the random arm.

Ideally, we would like to observe  $\mu_{ij}$  as the mean response for participants in preference group  $j$  who receive treatment  $i$ , for every  $(i, j)$  combination, but only four of the six possible subgroups are observable. We can observe participants in the choice arm who have a treatment preference, with expected outcomes  $\mu_{11}$  and  $\mu_{22}$  for participants who prefer and receive treatments A or B, respectively. Because undecided participants are re-randomised, we can estimate  $\mu_{13}$  and  $\mu_{23}$  from the choice arm. The remaining two subgroups, corresponding to  $\mu_{12}$  and  $\mu_{21}$ , are unobservable, because all participants in the choice arm with a treatment preference will receive that treatment.

Among the  $m$  participants in the choice group, there are  $m_1$  and  $m_2$  who choose treatments A and B, respectively. Among the remaining participants with no

preference, we assume that  $m_3/2$  are randomised to each of A and B. In the random arm, there is only one observed group for each treatment, containing unknown mixtures of treatment preferences. Here we observe samples of size  $n_1$  and  $n_2$  under treatments A and B.

Because of the potential effects of treatment preferences, we take a general approach and allow for different expected outcomes for the corresponding treatment groups in the choice and random arms, and for indifferent participants to have different expected outcomes from the other observable groups. We comment later on the implications of alternative assumptions concerning the responses in the indifferent participants.

We define the direct treatment effect by  $\Delta\tau = \tau_1 - \tau_2$ , or equivalently  $\mu_1 - \mu_2$ . One can estimate  $\Delta\tau$  from the random arm, through the sample estimate  $\bar{Y}_1 - \bar{Y}_2$  of  $\mu_1 - \mu_2$ . By virtue of the randomisation, this estimate will be unbiased, because participant preferences will be balanced in expectation. However, the estimate will be less precise than in a parallel group trial of the same total size.

We now consider the selection and preference effects. Because there are three preference groups, there are two possible contrasts for each effect.

**Selection effect**

The first selection effect is defined by  $\Delta v = v_1 - v_2$ , being the expected difference in responses between participants who would choose treatment A versus treatment B, if allowed to do so. Equivalently

$$\Delta v = [(\mu_{11} + \mu_{21}) - (\mu_{12} + \mu_{22})] / 2, \tag{2}$$

which involves two unobservable means,  $\mu_{12}$  and  $\mu_{21}$ . The first stage randomisation implies equal expected distributions of preferences in the choice and random arms. Hence, the mean for the random arm under treatment A is a weighted average of the means in the three preference groups; thus

**Table 1** Observations and parameters by group and actual treatment for the 2-stage randomised design

Actual treatment		Choice arm (Total sample size = m)			Random arm (Total sample size = n)
		Choose A	Choose B	Undecided (and randomised)	
A	Sample size	$m_1$	—	$m_3/2$	$n_1$
	Observations	$x_{11}, \dots, x_{1,m_1}$		$v_{11}, \dots, v_{1,m_3/2}$	$y_{11}, \dots, y_{1,n_1}$
	Mean	$\mu_{11}$	$(\mu_{12})$	$\mu_{13}$	$\mu_1$
	SD	$\sigma_{11}$		$\sigma_{13}$	$\sigma_1$
B	Sample size	—	$m_2$	$m_3/2$	$n_2$
	Observation		$x_{21}, \dots, x_{2,m_2}$	$v_{21}, \dots, v_{2,m_3/2}$	$y_{21}, \dots, y_{2,n_2}$
	Mean	$(\mu_{21})$	$\mu_{22}$	$\mu_{23}$	$\mu_2$
	SD		$\sigma_{22}$	$\sigma_{23}$	$\sigma_2$

$\mu_{12}$  and  $\mu_{21}$  are not observable

$$\mu_1 = \alpha\mu_{11} + \beta\mu_{12} + \gamma\mu_{13}$$

from which the unobservable  $\mu_{12}$  can be expressed in terms of observable quantities, as

$$\mu_{12} = [\mu_1 - \alpha\mu_{11} - \gamma\mu_{13}] / \beta,$$

and similarly for  $\mu_{21}$ . From (1) we can show that  $v_j = \frac{1}{2} [\mu_{1j} + \mu_{2j}] - \mu$ , for  $j = 1, 2$ . By substituting for the two unobservable means in (2), we derive  $\Delta v$  in terms of observables, as:

$$\Delta v = \frac{1}{2} \left[ \mu_{11} \left( 1 + \frac{\alpha}{\beta} \right) - \frac{\mu_1}{\beta} + \frac{\mu_2}{\alpha} - \mu_{22} \left( 1 + \frac{\beta}{\alpha} \right) - \frac{\mu_{23}\gamma}{\alpha} + \frac{\mu_{13}\gamma}{\beta} \right] \quad (3)$$

To estimate  $\Delta v$ , we may substitute into (3) the six observable sample means, and empirical estimates of the preference proportions from the choice arm,  $\alpha = m_1/m$ ,  $\beta = m_2/m$ , and  $\gamma = m_3/m$ . After simplification, we obtain

$$\hat{\Delta v} = [(z_1 - z_2) - \hat{\gamma}(w_1 - w_2)] / [2\hat{\alpha}\hat{\beta}m] \quad (4)$$

where  $z_1 = m_1(\bar{X}_1 - \bar{Y}_1)$ ,  $z_2 = m_2(\bar{X}_2 - \bar{Y}_2)$ ,  $w_1 = m_1(\bar{X}_1 - \bar{V}_1)$  and  $w_2 = m_2(\bar{X}_1 - \bar{V}_2)$ . If there are no indifferent participants ( $\gamma = 0$ , and hence  $\beta = 1 - \alpha$ ), result (4) simplifies to the same estimator as in our previous work ([3], Eq. 2).

**Preference effect**

The preference effect can be regarded as the interaction between a participant’s preferred treatment and the treatment actually received, or equivalently the difference between the treatment effects (A - B) for participants who would have chosen A or B. It can be defined as

$$\Delta\pi = (\pi_{11} - \pi_{12} - \pi_{21} + \pi_{22}) / 2$$

or in terms of the mean outcomes

$$\Delta\pi = [(\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22})] / 2$$

This involves the unobservables  $\mu_{12}$  and  $\mu_{21}$ , but if we again invoke the equivalence of the preference distributions in the choice and random arms, we obtain

$$\Delta\pi = \frac{1}{2} \left[ \mu_{11} \left( 1 + \frac{\alpha}{\beta} \right) - \frac{\mu_1}{\beta} - \frac{\mu_2}{\alpha} + \mu_{22} \left( 1 + \frac{\beta}{\alpha} \right) + \frac{\mu_{13}\gamma}{\beta} + \frac{\mu_{23}\gamma}{\alpha} \right]$$

which involves only estimable quantities. Similarly to the derivation for the selection effect, we estimate the preference effect as:

$$\hat{\Delta\pi} = [(z_1 + z_2) - \hat{\gamma}(w_1 + w_2)] / [2\hat{\alpha}\hat{\beta}m] \quad (5)$$

Again, this devolves to a simpler previous result when there are no undecided participants ([3], Eq. 3).

**Second contrasts for selection and preference effects**

For the selection effect, we can define a second contrast as:

$$\Delta v' = v_3 - \frac{v_1 + v_2}{2}$$

which represents the difference in outcomes between participants with no preference vs. those with a definite treatment preference. This is orthogonal to  $\Delta v$ , and can be expressed as

$$\Delta v' = \left[ (\mu_{13} + \mu_{23}) - \frac{1}{2} (\mu_{11} + \mu_{21} + \mu_{12} + \mu_{22}) \right] / 2$$

which is equivalent to

$$\Delta v' = \frac{1}{4\alpha\beta} [\alpha(\mu_{11} - \mu_1) + \beta(\mu_{22} - \mu_2) - 2\alpha\beta(\mu_{11} + \mu_{22} - \mu_{13} - \mu_{23}) - \gamma\{\alpha(\mu_{11} - \mu_{13}) + \beta(\mu_{22} - \mu_{23})\}]$$

when expressed in terms of estimable quantities. By again substituting empirical sample means and preference proportions, we have an estimator

$$\hat{\Delta v}' = [(z_1 + z_2) - (w_1 + w_2) + (\hat{\alpha} - \hat{\beta})(w_1 - w_2)] / [4\hat{\alpha}\hat{\beta}m].$$

Similarly, a suitable second contrast for the preference effect is

$$\Delta\pi' = \left[ \left( \frac{\pi_{11} + \pi_{12}}{2} - \pi_{13} \right) - \left( \frac{\pi_{21} + \pi_{22}}{2} - \pi_{23} \right) \right] / 2,$$

which represents the difference in treatment effects (A vs. B) between participants with no preference vs. those with any definite preference. It is orthogonal to  $\Delta\pi$ , and in terms of the observable means is

$$\Delta\pi' = \frac{1}{4\alpha\beta} [\alpha(\mu_1 - \mu_{11}) - \beta(\mu_2 - \mu_{22}) + 2\alpha\beta(\mu_{11} - \mu_{22} - \mu_{13} + \mu_{23}) + \gamma\{\alpha(\mu_{11} - \mu_{13}) - \beta(\mu_{22} - \mu_{23})\}]$$

with a corresponding estimator

$$\hat{\Delta\pi}' = [-(z_1 - z_2) + (w_1 - w_2) - (\hat{\alpha} - \hat{\beta})(w_1 + w_2)] / [4\hat{\alpha}\hat{\beta}m]$$

**Significance tests and confidence intervals for selection and preference effects**

To test the null hypothesis of no selection effect ( $H_0: \Delta v = 0$ ), we can use the inner component of (4) as a test statistic, i.e.,

$$T = (z_1 - z_2) - \hat{\gamma}(w_1 - w_2) \quad (6)$$

An approximately normal z-statistic is given by  $z = T / \text{var}(T)$ , where

$$\begin{aligned} \text{var}(T) = & m\{\alpha d_1^2 + \beta d_2^2 - (\alpha d_1 - \beta d_2)^2 + (1-\gamma)^2(\alpha\sigma_{11}^2 + \beta\sigma_{22}^2) \\ & + 2\left(\frac{\theta}{1-\theta}\right)(\alpha^2\sigma_1^2 + \beta^2\sigma_2^2) + 2\gamma(\alpha^2\sigma_{13}^2 + \beta^2\sigma_{23}^2) \\ & + \gamma(1-4\gamma)(\alpha e_1 - \beta e_2)^2 + \gamma^2(\alpha e_1^2 + \beta e_2^2) \\ & - 2\gamma[\alpha(1-2\alpha)d_1e_1 + \beta(1-2\beta)d_2e_2 + 2\alpha\beta(d_1e_2 - e_1d_2)]\} \end{aligned} \tag{7}$$

with  $d_1 = \mu_{11} - \mu_{1\cdot}$ ,  $d_2 = \mu_{22} - \mu_{2\cdot}$ ,  $e_1 = \mu_{11} - \mu_{13}$ ,  $e_2 = \mu_{22} - \mu_{23}$ ,  $\theta = m/N$  (the proportion of participants assigned to the choice arm), and where we have assumed for simplicity  $n_1 = n_2 = N/2$ . An outline of this and later derivations is given in the Additional file 1.

Because  $\hat{\Delta v} = T/[2\hat{\alpha}\hat{\beta}m]$ , further development gives

$$\begin{aligned} \text{var}(\hat{\Delta v}) \approx & \left[ \text{var}(T) + \frac{T^2}{\alpha^2\beta^2} \text{var}(\hat{\alpha}\hat{\beta}) \right. \\ & \left. - \frac{2T}{\alpha\beta} \text{cov}(T, \hat{\alpha}\hat{\beta}) \right] / [2\alpha\beta m]^2 \end{aligned} \tag{8}$$

where *cov* denotes a covariance. This leads to

$$\begin{aligned} \text{var}(\hat{\Delta v}) \approx & \left[ \text{var}(T) + \frac{T^2}{m\alpha\beta}(\alpha + \beta - 4\alpha\beta) \right. \\ & - \frac{2T\gamma}{\alpha\beta} \{ \alpha(1-2\alpha)d_1 - \beta(1-2\beta)d_2 \\ & \left. - \alpha\beta[(1-4\alpha)e_1 - (1-4\beta)e_2] \} \right] / [2\alpha\beta m]^2 \end{aligned} \tag{9}$$

We can use variance (9) to establish a confidence interval for  $\Delta v$ .

Similar results emerge for the preference effect  $\Delta\pi$ . For testing the null hypothesis of no preference effect ( $H_0: \Delta\pi = 0$ ) we use

$$T^* = (z_1 + z_2) - \hat{\gamma}(w_1 + w_2)$$

with

$$\begin{aligned} \text{var}(T^*) = & m\{\alpha d_1^2 + \beta d_2^2 - (\alpha d_1 + \beta d_2)^2 \\ & + (1-\gamma)^2(\alpha\sigma_{11}^2 + \beta\sigma_{22}^2) + 2\left(\frac{\theta}{1-\theta}\right)(\alpha^2\sigma_1^2 + \beta^2\sigma_2^2) \\ & + 2\gamma(\alpha^2\sigma_{13}^2 + \beta^2\sigma_{23}^2) + \gamma(1-4\gamma)(\alpha e_1 + \beta e_2)^2 \\ & + \gamma^2(\alpha e_1^2 + \beta e_2^2) - 2\gamma[\alpha(1-2\alpha)d_1e_1 + \beta(1-2\beta)d_2e_2 \\ & - 2\alpha\beta(d_1e_2 - e_1d_2)]\} \end{aligned} \tag{10}$$

This leads to a sample variance of the preference effect itself as

$$\begin{aligned} \text{var}(\hat{\Delta\pi}) \approx & \left[ \text{var}(T^*) + \frac{T^{*2}}{m\alpha\beta}(\alpha + \beta - 4\alpha\beta) \right. \\ & - \frac{2T^*\gamma}{\alpha\beta} \{ \alpha(1-2\alpha)d_1 - \beta(1-2\beta)d_2 \\ & \left. + \alpha\beta[(1-4\alpha)e_1 - (1-4\beta)e_2] \} \right] / [2\alpha\beta m]^2 \end{aligned} \tag{11}$$

Similar derivations are possible for testing and estimating the second contrasts  $\Delta v'$  and  $\Delta\pi'$ ; for brevity these are not shown here, but some simplified results are given below.

The results derived thus far are completely general in allowing for different variances in the various study subgroups, and for sample variation in the preference distribution within the choice arm. Considerable simplification occurs if these conditions are relaxed. First, we may assume that the outcome variance is constant. Second, we may ignore sample variability in the preference distribution in the choice arm. We refer to the variances as unconditional when sample variability in preferences is taken into account (as in Eqs. 7, 8, 9, 10 and 11), and conditional when the preference distribution is taken as fixed.

Making both assumptions together, the variances of the first estimated selection and preference contrasts become equal:

$$\begin{aligned} \text{var}(\hat{\Delta v}) = \text{var}(\hat{\Delta\pi}) = & \frac{\sigma^2}{4\alpha^2\beta^2 m} [(1-\gamma)^3 \\ & + 2(\alpha^2 + \beta^2)\left(\gamma + \frac{\theta}{1-\theta}\right)] \end{aligned} \tag{12}$$

as do the variances of the second estimated contrasts:

$$\begin{aligned} \text{var}(\hat{\Delta v}') = \text{var}(\hat{\Delta\pi}') = & \frac{\sigma^2}{16\alpha^2\beta^2\gamma m} [\gamma(1-\gamma)(\alpha-\beta)^2 \\ & + 2\{\alpha^2(2\beta + \gamma)^2 + \beta^2(2\alpha + \gamma)^2\} \\ & + 2\gamma(\alpha^2 + \beta^2)\left(\frac{\theta}{1-\theta}\right)] \end{aligned} \tag{13}$$

(detailed derivations not shown). As shown later in our example, the unconditional and conditional variances may be numerically very similar. Thus, in many practical situations, it will be sufficient to use the simpler conditional results (12) and (13).

#### Possible assumptions for responses in undecided participants

In the original development of the two-stage design [2], Rucker proposed that one should either assume that there were no undecided participants ( $\gamma = 0$ ), or that



these people were not subject to selection or preference effects ( $\mu_{i3} = \mu_i$  for  $i = 1, 2$ , respectively). Because the  $\{\mu_{i3}\}$  are estimable from undecided participants within the choice arm, one can test the null hypotheses  $H_0: \mu_{13} = \mu_1$  and  $H_0: \mu_{23} = \mu_2$  by comparing the outcomes of the undecideds to their similarly treated counterparts in the random arm. When testing selection and preference effects, these assumptions are equivalent to assuming that  $e_1 = e_2 = 0$ ; however, these terms appear in expressions such as (7) and (9, 10 and 11), implying that ignoring them can bias the test statistics or confidence intervals of these parameters.

An alternative approach is to assume that  $\mu_{ii} = \mu_{i3}$  for  $i = 1, 2$ , or that the expected outcomes are the same in participants with or without a preference, for each treatment. This amounts to the so-called “exclusion restriction” whereby the expected outcome is determined only by the treatment actually received. Again these assumptions are testable in the two-stage design, using data from the choice arm.

Another strategy would be to exclude undecided participants from the analysis. This approach would make it appear (incorrectly) that all participants in the random arm have a preferred treatment.

One can show that expressions (4) and (5) give consistent estimators of the selection and preference effects even in the presence of indifferent participants. Conversely if either set of assumptions above is adopted, but is actually false, biased estimation will occur. The bias will be important unless the proportion of indifferent participants is small.

**Results**

To illustrate these calculations, we will use data from a trial to compare alternative interventions for heavy menstrual bleeding [6, 7]. The treatment options were surgery or medical management without surgery. Women were randomised to a “conventional trial” subgroup (random arm) or to a “patient preference” subgroup (choice arm). In the choice arm, women were asked if they had a preference for one of the treatments, and if so, they could have it.

There were 227 women with outcome data available, of whom 130 (57%) were assigned to the choice arm. Within the choice arm, 19 elected to take medical treatment, and 21 decided to have surgery. The remaining 90 women (69%) had no preference, and were randomised (45 to each treatment). Thus while the rate of undecided participants was high, the two treatments were selected in somewhat similar numbers by women who were able to make a choice.

Table 2 summarises one principal study outcome, observed using a bleeding score (higher values represent poorer outcomes). There were 5 women in the choice arm and 41 in the random arm who refused to participate after entering the study. The disparity in these numbers has the potential to create biased comparisons among the women who did agree to participate, but unfortunately outcomes are unavailable for the refusers. The lower refusal rate in the choice arm is in the expected direction, because some participants will be attracted by possibly being able to choose their own treatment.

The investigators noted some baseline similarities between those choosing or being randomised to treatment, and they therefore aggregated the results from the randomised women in the random and choice arms, and reported outcomes only according to the actual treatment. In this paper we are able to use more detailed data on preferences within the choice arm (kindly provided by the investigators), allowing us to structure the results in the format of our Tables 1 and 2.

The direct effect of treatment (estimated from the random arm), was 12.10, with a standard error of 1.54 based on a standard deviation pooled over all the data. The 95% confidence interval was (9.1, 15.1), indicating a significant treatment effect in the direction of better outcomes for women having surgery.

The estimated first selection effect contrast  $\Delta \hat{\nu}$  was 3.03; using the unconditional approach and allowing for different subgroup variances, with standard error 6.72. Testing this effect gives  $T = 18.6$  ( $SD = 40.8$ );  $z = 0.46$  ( $p = 0.65$ , 2-sided test). The conditional standard error of  $\Delta \hat{\nu}$  when the outcome variance was taken as constant

**Table 2** Summary of results from a two-stage randomised trial of medical vs. surgical treatment

Actual treatment		Choice arm (Total sample size = m)			Random arm (Total sample size = n)
		Chose medical	Chose surgical	Undecided (and randomised)	
Medical	Sample size	19		45	49
	Mean	16.6		18.4	17.2
	SD	8.7		10.7	5.2
Surgical	Sample size		21	45	48
	Mean		5.9	4.3	5.1
	SD		7.2	5.2	7.7

across study groups was 6.64, only fractionally smaller than its unconditional counterpart. For simplicity, conditional standard errors were used in all further analyses.

The second selection contrast  $\Delta\hat{\nu}$  was 0.57, with conditional standard error 3.62. While there is no strong evidence of selection effects, the first contrast indicates better outcomes for women who would select surgery (treatment B) if allowed to do so, this being in addition to the direct benefit of surgery.

The two estimated preference effects are  $\Delta\hat{\pi} = 0.93$  and  $\Delta\hat{\pi} = -0.49$  (standard error = 6.64 for both). The first contrast, being positive (although not significant), indicates that the medical versus surgical effect is greater for women who preferred medical treatment.

We can also empirically test the alternative assumptions for the outcomes among the undecideds. First, using Rucker's approach [2], we can compare the outcomes for women on medical treatment, between the undecided participants and those in the random arm: this gives a mean difference of  $18.40 - 17.20 = 1.20$  ( $z = 0.77$ ,  $p = 0.44$ ). Similar analysis of the corresponding groups of women on surgery gives  $z = -0.51$  ( $p = 0.61$ ). Second, an assessment of the "exclusion restriction" assumption leads to similar findings, with only small differences in outcomes between women who chose their treatment in the choice arm, or were assigned to it in the random arm. Hence, in this example, there is no compelling evidence of systematically different outcomes for undecided women. However, we feel it is still advantageous to allow for the existence of undecided participants, in order to avoid bias. As noted earlier, we do not recommend simply ignoring the data from undecided participants.

## Discussion

We have extended the methodology for two-stage randomised trials to allow for participants who do not have a preferred treatment. As noted earlier, the two-stage design has been used in a variety of settings, and in some cases the selection and preference effects have been important, even when lacking a large treatment effect. From the perspective of shared decision-making concerning treatment options, it would be important to know about these effects.

Investigators who consider adopting the two-stage trial design would presumably hope that most participants have a preferred treatment. For instance, trials using patient decision aids try to improve patient understanding of the treatment options and reduce the number of patients who are uncertain about their treatment, to help them make an informed choice reflecting their own values [19]. For patients who are unable or unwilling to choose, the decision will likely revert to the clinician

proposing a treatment, and obtaining the tacit consent of the patient.

Despite the complication of undecided participants, the two-stage design remains attractive to people who do have a preferred treatment, and it still provides important information that is unobtainable from a conventional design. Preference effects may enhance the direct effect of treatment (as in our example), or act in the opposite direction. One may also compare the responses of decided vs. undecided individuals, as this may also be predictive of better or worse outcomes; these comparisons are accessible through our defined second contrasts of the selection and preference effects. Depending on the context, it seems better to allow study participants the option of declaring no preference, rather than pressuring them into a "forced" choice of treatment; in this scenario, "no preference" is a valid observation, and not "missing". However, the question of undecided patients may not arise in some circumstances, for instance where one treatment is usual care (e.g., [4, 5]).

We remarked earlier that it is inappropriate to simply ignore undecided participants in the choice arm, because of the potential for bias. Interestingly, in our example study [6, 7], the investigators elected to include the undecided participants, but they aggregated them with similarly treated participants in the random arm. Because all patients in this analysis are randomised, the estimated treatment effect is unbiased, but it lacks generalizability because participants with a declared preference are excluded.

Despite the advantages of the two-stage design, it must be recognised that the existence of undecided or indifferent patients may reduce its appeal. First, if a large proportion of participants cannot decide on treatment, estimating all the effects of interest (treatment, selection and preference effects) becomes less precise. Second, the presence of undecided participants leads to greater analytic complexity. Finally, the burden of explaining the study to participants may become unattractive if relatively few people will actually be able to make a treatment choice. However, the investigators in our example found that using this design "did not affect recruitment" compared to conventional approaches [6, 7]. Nevertheless, it should be noted that despite this encouraging finding, outcomes may still differ between participants in the random arm and the subgroup who are randomised within the choice arm.

In our example [6, 7], there was a high proportion of undecided participants, with less than one third of the patients in the choice arm declaring a preference. In this study, the information sheet for women in the choice arm contained only two additional sentences ("*We do realise that some women may have a preference for one of the treatments. If you feel strongly that you want one*

*treatment in particular, then please tell us*") compared to the information provided in the random arm. The mention of "strong" preferences may have deterred some women from choosing a treatment for which they had only a weaker preference. Researchers carrying out preference trials should consider carefully the methods they use to elicit preferences and their impact on the preferences that participants feel able to express.

Participants with no preference were also documented in another two-stage design [9], a trial comparing medication and cognitive behavioural therapy for the treatment of depression. Preferences were recorded on a 5-point scale, with participants (19%) who chose the middle value being defined as having "no clear treatment preference" or "indecisive". In other examples of the two-stage design [8, 16–18, 20], no undecided participants were reported, but it is unclear if such individuals were simply ignored, or if initially undecided persons were somehow persuaded to make a definitive choice before continuing into the study.

Further work is required to extend our earlier results [3] on optimising study efficiency while allowing for undecided participants. Optima here will depend on the proportion of participants who can decide on a treatment, as well as factors such as the preference rates for the two treatments among the decided individuals. The preference distribution will also affect the power to detect selection and preference effects [21].

We limited our attention to trials with equal numbers of participants being assigned to each treatment group in the random arm. Another research topic of future interest will therefore be to extend our methods to allow for unequal treatment group sizes. Unequal assignments might be desirable for ethical reasons, for instance if there is a large majority preference for one treatment, or if one treatment has a larger response variance. In such cases, the investigators might wish to improve study efficiency, by striving for overall balance in the trial, for example. We have discussed the ethics of unequal randomisation in somewhat more detail elsewhere [1].

We adopted a model in which preferences were assumed to have fixed effects. This was to conform to Rucker's original development, and indeed to our own earlier work in this area, which we have extended here. A further interesting extension would be to allow for preferences (and, potentially, the treatment) to have random effects. We note that all our estimates and tests are based on large sample normal approximations, and further work would be needed to consider small sample situations. One could, for instance, replace our z-statistics with t-statistics. Our previous work on sample size and power for the two-stage design [21] showed that particularly small studies will be inadequate to reliably estimate the parameters of interest. Rucker's investigation of

empirical type I error rates showed excellent adherence to nominal type I error rates with as few as 20 patients in each of the random and choice arms [2]. These results suggest that the large sample approximations described here will be adequate for most practical problems.

A reviewer pointed out that participants in the random arm may be randomised to a treatment that they do not like, and this in turn may influence outcomes. This is a valid point, but one that also affects a number of other designs, including standard parallel group trials, in exactly the same way. It has previously been noted that parallel group trials involve arms that may not be comparable after randomisation, because some patients will be pleased to get their preferred treatment, while others will be disappointed (this assuming that participants are not blinded) [22]. This phenomenon has been termed "reluctant acquiescence" [23], and it may in turn affect the estimated treatment effect. To the extent this is a problem for the two-stage design, exactly the same problem exists in parallel group designs, in terms of the impact on the estimated treatment effect. However, note that because the estimates of the selection and preference effects in a two-stage design are based partly on the data from the random arm, there may also be subsidiary implications for those estimates as well.

There are related issues in a variety of other preference-based designs, as we have reviewed elsewhere [1]. For example, there are challenges in how to present information about treatments and then eliciting preferences reliably, and one must recognise that there are potential differences between what are stated as the preferences for treatment in a general sense, and actual treatments chosen when participants are allowed to choose.

As we have reviewed elsewhere [1], randomised trial designs can allow for patient preference effects in various ways. There is also some literature on the impact of clinician preferences for treatments, and the associated "expertise bias" [24, 25]. The various designs that accommodate patient or clinician preferences are differentiated by whose preference is ascertained, how and when they are measured, and if treatment assignments take those preferences into account. Among these designs, the two-stage approach is unique in being able to estimate selection and preference effects in addition to the direct treatment effects [26, 27]. It remains a candidate design even when the complication of indifferent participants arises. The results in this paper have shown how the information from such participants can be used in the analysis for two-stage trials, to further elucidate the impact of treatment preferences on study outcomes.

## Conclusions

Treatment preferences among clinical trial participants are potentially important determinants of study outcomes,



and we have demonstrated how data from two-stage randomised trials can be analysed to estimate such effects. We have extended our earlier work to now allow for the possibility of participants who are indifferent or undecided about a preferred treatment, and we discussed alternative possible assumptions about their response patterns, compared to participants with a declared treatment preference. We also have shown how to obtain unbiased estimates and tests of the preference effects in the trial as a whole. Despite the presence of indifferent or undecided participants, the two stage design is nevertheless an attractive option for the examination of preference effects, a feature which is not possible with conventional trial designs.

## Additional file

**Additional file 1:** The Appendix provides a more detailed outline of the derivations of the test statistics and variances associated with estimates of selection and preference effects described in this paper. (DOCX 64 kb)

### Abbreviations

SD: Standard deviation

### Acknowledgements

We thank Dr KG Cooper for provision of data from his study, used in our illustrative example.

### Funding

This work was partially supported by funds from the Natural Sciences and Engineering Research Council.

### Availability of data and materials

All data generated or analysed during this study are included within this published article. The study cited in references 6 and 7 provides the basis for our illustrative example. Dr Kevin Cooper (Consultant Gynaecologist, Aberdeen Royal Infirmary, Scotland) kindly provided additional details of the study that were not shown in those references, and gave us permission to use them in our published work. Specifically, these details were the provision of summary statistics (sample sizes, means, and standard deviations) of the study outcome variable, according to the randomly assigned arm of the study, and according to the stated treatment preferences in the choice arm. These summary statistics are shown within the current paper.

### Authors' contributions

All authors read and approved the final version of the manuscript. Their specific contributions are as follows: SW: provided the overall leadership for this work, and was primarily responsible for developing the analytic method and for data analysis; wrote the initial and final drafts of the paper. RT: contributed to the conception and design of the study, revised the manuscript for important intellectual content, and agrees to be accountable for all aspects of the work. PM: made substantial contributions to the conception and design of the study, particularly interpretation and derivation of equations; was involved in revising the manuscript for important intellectual content, and takes public responsibility for the content; agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. KM: commented on multiple drafts of the paper and made intellectual contribution to its content, and agrees to be accountable for the work. LI: was involved in revising the manuscript for important intellectual content; takes public responsibility for the content; and agrees to be accountable for the work.

### Authors' information

None.

### Competing interests

The authors declare that they have no competing interests.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

Not applicable.

### Author details

<sup>1</sup>Department of Clinical Epidemiology and Biostatistics, McMaster University, CRL 233, Hamilton, ON, Canada L8N 3Z5. <sup>2</sup>School of Public Health and Community Medicine, University of New South Wales, Sydney, NSW 2052, Australia. <sup>3</sup>Screening and Test Evaluation Program, Sydney School of Public Health, Sydney Medical School, University of Sydney, Sydney NSW 2006, Australia.

Received: 16 August 2016 Accepted: 2 February 2017

Published online: 20 February 2017

### References

- Walter SD, Turner R, Macaskill P, McCaffery K, Irwig L. Beyond the treatment effect: evaluating the effects of patient preferences in randomised trials. *Stat Methods Med Res.* 2014. doi:10.1177/0962280214550516.
- Rucker G. A two-stage trial design for testing treatment, self-selection and treatment preference effects. *Stat Med.* 1989;8:477–85.
- Walter SD, Turner RM, Macaskill P, McCaffery K, Irwig L. Optimal allocation of participants for the estimation of selection, preference and treatment effects in the two-stage randomised trial design. *Stat Med.* 2012;31:1307–22.
- McCaffery KJ, Turner R, Macaskill P, Walter SD, Chan SF, Irwig L. Determining the impact of informed choice: separating treatment effects from the effects of choice and selection in randomised trials. *Med Decis Making.* 2011;31:229–36.
- McCaffery KJ, Irwig L, Turner R, Chan SF, Macaskill P, Lewicka M, Clarke J, Weisberg E, Barratt A. Psychosocial outcomes of three triage methods for the management of borderline abnormal cervical smears: an open randomised trial. *BMJ.* 2010;340:b4491.
- Cooper KG, Grant AM, Garratt AM. The impact of using a partially randomised patient preference design when evaluating alternative managements for heavy menstrual bleeding. *Br J Obstet Gynaecol.* 1997; 104:1367–73.
- Cooper KG, Parkin DE, Garratt AM, Grant AM. Two-year follow up of women randomised to medical management or transcervical resection of the endometrium for heavy menstrual loss: clinical and quality of life outcomes. *Br J Obstet Gynaecol.* 1999;106:258–65.
- Floyd AHL, Moyer A. Effects of participant preferences in unblinded randomized controlled trials. *J Empir Res Hum Res Ethics.* 2010;8:81–93.
- Mergl R, Henkel V, Allgaier AK, Kramer D, Hautzinger M, Kohnen R, Coyne J, Hegerl U. Are treatment preferences relevant in response to serotonergic antidepressants and cognitive-behavioural therapy in depressed primary care patients? Results from a randomized controlled trial including a patients' choice arm. *Psychother Psychosom.* 2011;80:39–47.
- Rokke PD, Tomhave JA, Jovic Z. The role of client choice and target selection in self-management therapy for depression in older adults. *Psychol Aging.* 1999;14:155–69.
- Long Q, Little RJ, Lin X. Causal inference in hybrid intervention trials involving treatment choice. *J Am Stat Assoc.* 2008;103:474–84.
- Janevic MR, Janz NK, Dodge JA, Lin X, Pan W, Sinco BR, Clark NM. The role of choice in health education intervention trials: a review and case study. *Soc Sci Med.* 2003;56:1581–94.
- Rose JP, Geers AL, Rasinski HM, Fowler SL. Choice and placebo expectation effects in the context of pain analgesia. *J Behav Med.* 2012;35:462–70.
- Kernahan C, Bettencourt BA. Motivated processing of black and white targets: the situation of choice versus assignment. *Basic Appl Soc Psych.* 2002;24:125–44.
- Noel PH, Larme AC, Meyer J, Marsh G, Correa A, Pugh JA. Patient choice in diabetes education curriculum. Nutritional versus standard content for type 2 diabetes. *Diabetes Care.* 1998;21:896–901.
- Burke LE, Styn MA, Steenkiste AR, Music E, Warziski M, Choo J. A randomized clinical trial testing treatment preference and two dietary

- options in behavioural weight management: preliminary results of the impact of diet at 6 months - PREFER study. *Obesity*. 2006;14:2007–17.
17. Burke LE, Hudson AG, Warziski MT, Styn MA, Music E, Elci OU, Sereika SM. Effects of a vegetarian diet and treatment preference on biochemical and dietary variables in overweight and obese adults: a randomized clinical trial. *Am J Clin Nutr*. 2007;86:588–96.
  18. Burke LE, Warziski MT, Styn MA, Music E, Hudson AG, Sereika SM. A randomized clinical trial of a standard versus vegetarian diet for weight loss: the impact of treatment preference. *Int J Obes (Lond)*. 2008;32:166–76.
  19. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Col NF, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev*. 2014. doi:10.1002/14651858.
  20. Eisenberg DM, Post DE, Davis RB, Connelly MT, Legedza AT, Hrbek A, et al. Addition of choice of complementary therapies to usual care for acute low back pain - a randomized controlled trial. *Spine*. 2007;32:151–8.
  21. Turner RM, Walter SD, Macaskill P, McCaffery KJ, Irwig L. Sample size and power when designing a randomised trial for the estimation of treatment, selection and preference effects. *Med Decis Making*. 2014;34:711–9.
  22. McPherson K, Chalmers I. Incorporating patient preferences into clinical trials. *BMJ*. 1998;317:78.
  23. Tilbrook H. Patients' preferences within randomised trials: systematic review and patient-level meta-analysis. *BMJ*. 2008;337:a1864.
  24. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S, et al. Need for expertise-based randomized controlled trials. *BMJ*. 2005;330:88–91.
  25. Walter SD, Ismaila AS, Devereaux PJ, for the SPRINT study investigators (2008). Statistical issues in the design and analysis of expertise-based randomized clinical trials. *Stat Med*. 2008;27:6583–96.
  26. Goodman LA. On the exact variance of products. *JASA*. 1960;55:708–13.
  27. Johnson NL, Kotz S, Balakrishnan N. *Discrete multivariate distributions*. New York: John Wiley & Sons; 1997.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

