Contemporary Clinical Trials Communications 3 (2016) 117–121

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Random guess and wishful thinking are the best blinding scenarios

Heejung Bang

Division of Biostatistics, Department of Public Health Sciences, University of California, Davis, CA, USA

ARTICLE INFO

Article history:
Received 22 December 2015
Received in revised form
23 March 2016
Accepted 3 May 2016
Available online 7 May 2016

Keywords:
Bias
Blinding
Blinding index
Clinical trial

ABSTRACT

Blinding is a methodologic safeguard of treatment evaluation, yet severely understudied empirically. Mathieu et al.'s theoretical analysis (2014) provided an important message that blinding cannot eliminate potential for bias associated with belief about allocation in randomized controlled trial; just like the intent-to-treat principle does not guarantee unbiased estimation under noncompliance, the blinded randomized trial as a golden standard may produce bias. They showed possible biases but did not assess how large the bias could be in different scenarios. In this paper, we examined their findings, and numerically assessed and compared the bias in treatment effect parameters by simulation under frequently encountered blinding scenarios, aiming to identify the most ideal blinding scenarios in practice. We conclude that Random Guess and Wishful Thinking (e.g., participants tend to believe they received treatment) are the most ideal blinding scenarios, incurring minimal bias. We also find some evidence that imperfect or partial blinding can be better than no blinding.

1. Introduction

Blinding is a critical feature in comparative evaluations to minimize various biases. Blinded randomized controlled trial (RCT) is widely accepted as a gold standard when we compare treatments, whenever feasible [1–5]. Blinding can be more relevant and important to subjective outcomes, such as patient-reported outcomes. Several authors reviewed the current practice of blinding-related techniques, assessment and reporting [2,6], and it has been suggested that unblinding may overestimate the treatment effect [7,8]. Although the role and importance of blinding are well recognized in the clinical trial community, statistical investigation on this topic has been rare, partly due to inherently complicated and subjective/qualitative nature.

Recently, Mathieu et al. (2014) provided a theoretical analysis demonstrating that blinding cannot eliminate potential for bias associated with belief about allocation in RCT, which could be surprising or counterintuitive for many trialists [1,9]. Specifically, they studied a mathematical framework of simple RCTs, and identified conditions where the bias in treatment effect is equal to zero. Except for highly restrictive conditions, if belief about the treatment allocation is translated into the study outcome (e.g., over or under-reporting of the outcome), the bias is expected to be non-zero. Thus, the authors concluded that blinding cannot guarantee to prevent bias caused by belief, but emphasized that it is not their intention to suggest that RCTs should not be blinded. They considered deterministic/hypothetical scenarios under a type of effective blinding, without numerical evaluation.

In this paper, we intend to study Mathieu et al.'s findings carefully and assess the bias in different treatment effect parameters numerically in more practical/realistic settings, under qualitatively different blinding scenarios, with a goal to provide a better insight and some actionable advice for trialists, if any. In section 2, we present background and mathematical framework. In section 3, we perform simulation studies and summarize the findings. Discussions and conclusions are provided at the end.

2. Mathematical framework

We summarize a simple, theoretical framework posited by Mathieu et al. as basis, examine and adapt here [9]. Each cell in a 3×2 table for guess status by allocation has the number of subjects n_{ij}, where j denotes allocation (T = treatment, C = control) and i denotes belief about allocation (t = treatment, u = don't know, c = control); see Table 1. We assume that outcomes can be distorted via two mechanisms, where a_i is the magnitude of distortion that is independent of the true outcome Y_{ij} (e.g., fixed, *a priori* expectation before allocation) and b_i is that proportional to the outcome (e.g., unblinding during trial) in both arms. Thus, the total distortion or bias in the individual outcome due to belief in a cell is $a_i + b_i Y_{ij}$; the observed outcome is

E-mail address: hbang@ucdavis.edu.

http://dx.doi.org/10.1016/j.conctc.2016.05.003

Table 1
Statistical notation.

Treatment	Allocation (j), outcome & sample size			
Guess (i)	Treatment arm (T)		Control arm (C)	
Received treatment (t)	$Y'_{tT} = Y_{tT} + a_t + b_t Y_{tT}$	n_{tT}	$Y'_{tC} = Y_{tC} + a_t + b_t Y_{tC}$	n_{tC}
Don't know (u)	$Y'_{uT} = Y_{uT} + a_u + b_u Y_{uT}$	n_{uT}	$Y'_{uC} = Y_{uC} + a_u + b_u Y_{uC}$	n_{uC}
Received control (c)	$Y'_{cT} = Y_{cT} + a_c + b_c Y_{cT}$	n_{cT}	$Y'_{cC} = Y_{cC} + a_c + b_c Y_{cC}$	n_{cC}

See Table 1 in Mathieu et al. (2014).

$$Y'_{ij} = Y_{ij} + a_i + b_i Y_{ij}.$$

If there is no distortion, then $a_i = b_i = 0$ so $Y'_{ij} = Y_{ij}$, i.e., individual outcome is unbiased.

We consider three treatment effect parameters: mean difference, mean ratio and odds ratio (OR) where OR is relevant when the outcome is binary. For binary outcome, mean difference and ratio are typically interpreted as incidence or risk difference (RD) and risk ratio (RR) in epidemiologic and clinical trial contexts.

Let us assume the randomization of the 1:1 allocation ratio for two treatments, and define the n_{ij}-weighted mean of all Y'_{ij} as $\sum_i (n_{ij} Y'_{ij}) / \sum_i n_{ij}$ for j = T and C. If belief about allocation is independent of actual allocation, so $n_{iT} = n_{iC} = n_i$, which can be a specific form of effective blinding (see its connection to the 'Wishful thinking' scenario below) [9], then the absolute estimate of the treatment effect is unbiased:

$$\sum_i (n_{iT} Y'_{iT}) \Big/ \sum_i n_{iT} - \sum_i (n_{iC} Y'_{iC}) \Big/ \sum_i n_{iC} = \sum_i (n_{iT} Y_{iT}) \Big/ \sum_i n_{iT} - \sum_i (n_{iC} Y_{iC}) \Big/ \sum_i n_{iC}$$

if, but not only if, in every stratum, either (1) $b_i = 0$ or (2) $Y_{iC} = Y_{iT}$ (i.e., there is no effect of treatment). We can further show that when $n_{iT} = n_{iC} = n_i$ may or may not be true, if (3) $\sum_i n_{iT} a_i = 0$, $\sum_i n_{iT} b_i Y_{iT} = 0$, $\sum_i n_{iC} a_i = 0$, and $\sum_i n_{iC} b_i Y_{iC} = 0$, then unbiased estimation is achieved. Here, the conditions in (3) hold, for example, if $n_{tT} = n_{cT}$, $n_{tC} = n_{cC}$, $a_t = -a_c$, $b_t = -b_c$, $a_u = 0$, $b_u = 0$, $Y_{tT} = Y_{cT}$ and $Y_{tC} = Y_{cC}$, where this scenario can be realized when underlying true means are independent of guess (so that biases are introduced only via a's and b's); no bias among subjects who answered "Don't know"; and among those who provided treatment guesses, biases (due to over vs. under-reporting) cancel out *within* each arm. This situation can be regarded as another plausible form of effective blinding which yields a combination of various balances within each arm; see the 'Random guess' scenario below.

Next, under $n_{iT} = n_{iC} = n_i$, the relative estimate of the treatment effect is unbiased:

$$\left[\sum_i (n_{iT} Y'_{iT}) \Big/ \sum_i n_{iT}\right] \div \left[\sum_i (n_{iC} Y'_{iC}) \Big/ \sum_i n_{iC}\right] = \left[\sum_i (n_{iT} Y_{iT}) \Big/ \sum_i n_{iT}\right] \div \left[\sum_i (n_{iC} Y_{iC}) \Big/ \sum_i n_{iC}\right]$$

if, but not only if, in every stratum, either (1) $a_i = 0$ and (1a) all b_i's are equal or (1b) all Y_{iC}'s are equal and all Y_{iT}'s are equal; or (2) $Y_{iC} = Y_{iT}$. Moreover, when $n_{iT} = n_{iC} = n_i$ may or may not be true, if the same 4 conditions as in (3) above are met, unbiased estimation is achieved. Finally, the unbiasedness of OR is even rarer, with equality holding if, but not only if, either (1) $Y_{iC} = Y_{iT}$ or (2) $a_i = 0$ and $b_i = 0$, but unlikely otherwise. Yet, under some conditions, say, the rare disease assumption (e.g., <10%), RR and OR would be close [10,11].

Now, we introduce different blinding scenarios through representative guess status in (Treatment, Control) = (random, random), (correct, opposite), (correct, random), and (correct, correct), where we will call these four classifications 'Random guess', 'Wishful thinking', 'Unblinded in one arm', and 'Unblinded in both arms', respectively, for convenience. Nine blinding scenarios may be classified based on the proportion of correct guesses, and these four scenarios have been shown to be relatively common in systematic reviews [12–14] and are covered in this study; extensions to the remaining five scenarios are straightforward. For example, 'Random guess' may correspond mathematically to blinding index (BI) values of (0, 0), 'Wishful thinking' to (k, −k), 'Unblinded in one arm' to (k, 0), and 'Unblinded in both arms' to (k, k), with a positive proportion of k, where BIs for treatment arm and control arm are defined as:

$$BI_T = (2{*}n_{tT}/(n_{tT} + n_{cT}) - 1){*}(n_{tT} + n_{cT})/(n_{tT} + n_{uT} + n_{cT}),$$

$$BI_C = (2{*}n_{cC}/(n_{tC} + n_{cC}) - 1){*}(n_{tC} + n_{cC})/(n_{tC} + n_{uC} + n_{cC}).$$

BI may serve as an indicator of potential unblinding through quantifying 'imbalance' between the two statuses of the identified guesses, i.e., T vs. C [15]. Here, we chose Bang et al.'s BI because it is widely used in practice, including in meta-analyses, and it assesses blinding separately for different arms (unlike James et al.'s BI that provides one value) so that it could capture different blinding patterns in different arms [12,–14,16–19]. Roughly speaking, BI = 0 means that the proportions of correct and incorrect guesses are equal, adjusting for the count of "Don't know". Here, k = 20% has been used as an ad-hoc threshold for classification purposes [14,15,20]. Note that the condition, $n_{iT} = n_{iC} = n_i$, imposed in Mathieu et al. implies $BI_T = -BI_C$. Hence, in practice, if $BI_T \approx -BI_C \gg 0$, say, >20%, we may designate as the 'Wishful thinking' scenario. In contrast, if a set of the conditions in (3) above are satisfied (e.g., $n_{tT} = n_{cT}$, $n_{tC} = n_{cC}$), so $BI_T \approx -BI_C \approx 0$, we may designate as the 'Random guess' scenario. These two scenarios may constitute effective blinding.

Although these theoretical conditions identified in simplistic models/settings are critical in the improvement of our understanding about blinding and its potential impact on treatment effect, it is not straightforward to understand the extent of bias in any given trial. Exact cancellations would be nearly impossible and some conditions are too restrictive or highly implausible (for example, $b_i = b$), so that bias is highly likely in most cases, especially, in more general or realistic settings we simulated below.

3. Simulation study

3.1. Configuration and data generation

In this section, we examine the empirical bias in the treatment effect by simulation in the combinations of: 1) outcomes (continuous and binary); 2) parameters (mean difference/RD, mean ratio/RR, and OR); 3) hypotheses (null and non-null effects); and 4) different blinding scenarios. In brief, we followed Mathieu et al.'s