Letter to the Editor

Statistics in Medicine

Accepted draft: http://madeyski.e-informatyka.pl/download/KitchenhamMadeyskiCurtinSIM.pdf of the paper: Barbara Kitchenham, Lech Madeyski, and Francois Curtin, "Corrections to effect size variances for continuous outcomes of cross-over clinical trials", Statistics in Medicine, vol. 37, no. 2, pp. 320-323, 2018. DOI: 10.1002/sim.7379 published by Wiley and available at http://doi.org/10.1002/sim.7379

(www.interscience.wiley.com) DOI: 10.1002/sim.7379

Corrections to effect size variances for continuous outcomes of cross-over clinical trials

Barbara Kitchenham[†] and Lech Madeyski[*‡] and François Curtin[§]

Keywords: cross-over trials; continuous data; effect sizes; effect size variances; corrections

1. Introduction

We would like to make some corrections to the formulas presented in the 2002 *Statistics in Medicine* article by Curtin et al. [1]. That article presented formulas for the variances of standardized weighted mean difference of an AB/BA cross-over trial that would be comparable both with parallel designs and cross-over designs.

There are three main issues in the Curtin et al.'s paper [1] that we address in this communication. Firstly, the paper proposes a standardized effect size for cross-over trials that is inconsistent with the standardized effect sizes used for other repeated measures designs such as the pretest-posttest studies used in educational studies, see [2] and [3]. Secondly, the change to the standardized effect size for cross-over studies necessitates a change to variance of the standardized effect size. Thirdly, the variance of the standardized effect size comparable with parallel trials was not based on the distribution of a valid t-variable, so includes some errors. We follow Curtin et al.'s approach and base our revised variance equations on the moments of the non-central t-distribution, replacing the t-variable with a variable based on the effect size.

2. The standardized effect sizes from AB/BA cross-over trials

Using Curtin et al.'s notation, the original paper derived the expectation and variance of the standardized mean difference of the mean cross-over trial with a period effect from the equation:

$$\frac{\bar{d}_{XO}}{s_x} = \frac{\frac{1}{2}(\bar{d}_{AB} + \bar{d}_{BA})}{s_x} \tag{1}$$

$\bar{d}_{AB}$ and $\bar{d}_{BA}$ are the mean cross-over differences in sequences AB and BA respectively and s_x^2 is the pooled within sequence cross-over difference variance.

However, if we want a standardized effect size for cross-over designs that is consistent with other repeated measures, the unstandardized effect size $\bar{d}_{XO}$ should be standardized by the within-subject standard deviation, s_e [3]. s_e is also a natural choice for the standardizer because s_e^2 is the random effects residual term obtained when analysing cross-over data with a linear mixed model. Since Curtin et al. note that $s_x^2 = 2s_e^2$, it is easy to calculate $\frac{\bar{d}_{XO}}{s_e}$. For consistency with [3], we refer to this standardized effect size as d_{RM}, which is an estimate of the parameter δ_{RM}.

[*] *Correspondence to: Lech Madeyski, Faculty of Computer Science and Management, Wroclaw University of Science and Technology, Wyb.Wyspianskiego 27, 50-370 Wroclaw, Poland. Email: Lech.Madeyski@pwr.edu.pl*

[†] School of Computing and Mathematics, Keele University, Keele, Staffordshire ST5 5BG, UK.

[‡] Faculty of Computer Science and Management, Wroclaw University of Science and Technology, Wyb.Wyspianskiego 27, 50-370 Wroclaw, Poland.

[§] Research Center for Statistics, Geneva School of Economics and Management, University of Geneva, Switzerland and Geneuro SA, Geneva, Switzerland

Curtin et al. correctly propose basing a standardized effect size comparable with parallel trials on the within- plus between-subject variance, σ_b^2. For comparison with Morris and DeShon [3] we refer to this as d_{IG}, where IG refers to independent groups, so:

$$d_{IG} = \frac{\bar{d}_{XO}}{s_b} \tag{2}$$

while d_{IG} estimates the parameter δ_{IG}. Curtin et al. point out that $\sigma_b^2 = \sigma_\xi^2 + \sigma_e^2$ where σ_ξ^2 is the between-subject variance.

Morris and DeShon [3] point out that d_{RM} estimates the standardized expected change for individuals while d_{IG} estimates the standardized expected difference between the two methods. They note that either viewpoint might be the objective of meta-analysis.

3. Standardized effect size variances

The variance estimate most suitable for small samples (up to ≈ 30 participants) for any standardized mean difference effect size is derived from the non-central t distribution [3, 2]. Furthermore, Johnson and Welch [4] report the variance of a t variable with mean θ to be:

$$V(\theta) = \frac{df}{df-2}\left(1+\theta^2\right) - \frac{\theta^2}{\left[c(df)\right]^2} \tag{3}$$

Where θ is estimated by the sample t-value and $df = (n_{AB} + n_{BA} - 2)$ is the degrees of freedom associated with the t-test. Hedges [5] provides exact values of $c(df)$ for values of df up to 50. In addition, Morris [2] confirmed that $c(df)$ is well-approximated by $\left(1 - \frac{1}{4df-1}\right)$ even for small samples when estimating the variance of pretest-posttest standardized effect sizes.

If we know the relationship between θ and a standardized effect size δ is given by the equation:

$$\theta = A \times \delta \tag{4}$$

where A is a constant term, then, the variance of δ is:

$$var(\delta) = \frac{1}{A^2}V(\theta). \tag{5}$$

This is true for *any* standardized effect size that can be calculated from a t-value, including those obtained from repeated measures crossover designs, repeated measures pretest-posttest designs, and independent group designs.

In the case of a cross-over design, Senn [6] points out that the t test is based on:

$$t = \frac{2\bar{d}_{XO}}{s_x\sqrt{\left(\frac{1}{n_{AB}} + \frac{1}{n_{BA}}\right)}} \tag{6}$$

which suggests that s_x is a natural standardizer of *twice* the effect size. Furthermore, since $s_x = s_e\sqrt{2}$ and $\left(\frac{1}{n_{AB}} + \frac{1}{n_{BA}}\right) = \frac{(n_{AB}+n_{BA})}{n_{AB}n_{BA}}$

$$t = d_{RM}\sqrt{\frac{2n_{AB}n_{BA}}{(n_{AB}+n_{BA})}} \tag{7}$$

Thus,

$$var(\delta_{RM}) = V(\theta)\frac{(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} \tag{8}$$

Replacing θ by $\delta_{RM}\sqrt{\frac{2n_{AB}n_{BA}}{(n_{AB}+n_{BA})}}$ and employing Equation 3:

$$var(\delta_{RM}) = \left(\frac{df}{df-2}\right)\left[\frac{(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + \delta_{RM}^2\right] - \frac{\delta_{RM}^2}{\left[c(df)\right]^2} \tag{9}$$

This is similar to the equation for the variance of $\frac{\bar{d}_{XO}}{s_x}$ given in the Appendix to [1]. The difference is only in the term $2n_{AB}n_{BA}$ where Curtin et al. use the constant 4, and we use 2, because we standardize by s_e. However, for small sample sizes, it is inappropriate to replace δ_{RM} by d_{RM} in Equation 9, because d_{RM} is biased. For an unbiased estimate of δ_{RM}, we need to use the bias corrected estimate $g_{RM} = c(df) \times d_{RM}$ giving:

$$var(d_{RM}) = \left(\frac{df}{df-2}\right)\left[\frac{(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + g_{RM}^2\right] - \frac{g_{RM}^2}{[c(df)]^2} \quad (10)$$

In addition, multiplying the variance of d_{RM} by $c(df)^2$, an unbiased estimate of the variance of g_{RM} is:

$$var(g_{RM}) = c(df)^2\left(\frac{df}{df-2}\right)\left[\frac{(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + g_{RM}^2\right] - g_{RM}^2 \quad (11)$$

To construct the variance of d_{IG} it is necessary to consider the relationship between s_e and s_b. This relies on the correlation ρ between values obtained from the same subject:

$$\rho = \frac{\sigma_\xi^2}{\sigma_b^2} = \frac{\sigma_b^2 - \sigma_e^2}{\sigma_b^2} \quad (12)$$

so, $s_e = s_b\sqrt{(1-\hat{\rho})}$. Thus, based on Equation 1 and Equation 6, the relationship between t and d_{IG} is:

$$t = d_{IG}\sqrt{\frac{2n_{AB}n_{BA}}{(1-\hat{\rho})(n_{AB}+n_{BA})}} \quad (13)$$

Therefore replacing δ_{RM} with $\frac{\delta_{IG}}{\sqrt{(1-\hat{\rho})}}$ in Equation 9 and multiplying by $(1-\hat{\rho})$ gives:

$$var(\delta_{IG}) = \left(\frac{df}{df-2}\right)\left[\frac{(1-\hat{\rho})(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + \delta_{IG}^2\right] - \frac{\delta_{IG}^2}{[c(df)]^2} \quad (14)$$

Compared with the Equation 12 in [1], Equation 14 includes the term $(1-\hat{\rho})$ in the first term on the right-hand side of the equation and the term $2n_{AB}n_{BA}$ rather than $4n_{AB}n_{BA}$. The inclusion of the term $(1-\hat{\rho})$ in Equation 14 is comparable with the equivalent equation for the variance of pretest-posttest standardized effect size [2].

However, again, if we want the most appropriate variance for small sample sizes, we should not replace δ_{IG} by d_{IG}. Like d_{RM}, d_{IG} is biased for small sample sizes, so we need to replace δ_{IG} with $g_{IG} = c(df) \times d_{IG}$ giving:

$$var(d_{IG}) = \left(\frac{df}{df-2}\right)\left[\frac{(1-\hat{\rho})(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + g_{IG}^2\right] - \frac{g_{IG}^2}{[c(df)]^2} \quad (15)$$

In addition, the variance of g_{IG} is:

$$var(g_{IG}) = c(df)^2\left(\frac{df}{df-2}\right)\left[\frac{(1-\hat{\rho})(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + g_{IG}^2\right] - g_{IG}^2 \quad (16)$$

4. Approximate standardized effect sizes for larger samples

For an approximate standardized effect size, Curtin et al. [1] use the formula proposed by [5] for $var(d_{RM})$. Since the approximation assumes large sample sizes, the effect of the small sample size adjustment is negligible. So, after correcting a typographical error in [1], the approximate variance for d_{RM} is

$$var(d_{RM})_{Approx} = \frac{(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + \frac{d_{RM}^2}{2(n_{AB}+n_{BA}-3.94)} \quad (17)$$

In addition, based on the relationship between d_{RM} and d_{IG}, the large sample size approximation of the variance of d_{IG} is:

$$var(d_{IG})_{Approx} = (1-\hat{\rho})\frac{(n_{AB}+n_{BA})}{2n_{AB}n_{BA}} + \frac{d_{IG}^2}{2(n_{AB}+n_{BA}-3.94)} \quad (18)$$

References

1. Curtin F, Altman DG, Elbourne D. Meta-analysis combining parallel and cross-over clinical trials. I: Continuous outcomes. *Statistics in Medicine* 2002; **21**:2132–2144, doi:10.1002/sim.1205.
2. Morris SB. Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology* 2000; **53**:17–29.
3. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods* 2002; **7**(1):105–125, doi:10.1037//1082-989X.7.1.105.
4. Johnson NL, Welch BL. Applications of the non-central t-distribution. *Biometrika* 1940; **31**(3/4):362–389.
5. Hedges LV, Olkin I. *Statistical methods for meta-analysis*. Academic Press: Orlando, Florida, USA, 1985.
6. Senn S. *Cross-over Trials in Clinical Research*. 2nd edn., John Wiley and Sons, Ltd., 2002.