Revisiting and expanding the meta-analysis of variation: The log coefficient of variation ratio, lnCVR

Alistair M. Senior[1,2,*], Wolfgang Viechtbauer[3], and Shinichi Nakagawa[4,5,*]

1. University of Sydney, Charles Perkins Centre, Sydney, NSW 2006, Australia.

2. University of Sydney, School of Life and Environmental Sciences, Sydney, NSW 2006, Australia.

3. Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Faculty of Health, Medicine, and Life Sciences, Maastricht University, 6200 MD Maastricht, The Netherlands

4. Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia.

5. Diabetes and Metabolism Division, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, Sydney, NSW 2010, Australia.

* Authors to whom correspondence ought to be addressed.

Email: alistair.senior@sydney.edu.au

Email: s.nakagawa@unsw.edu.au

Tel: +64 (0) 286 270 703

Running Head: Meta-Analysis of lnCVR

Keywords:

Cross-Over Design, Cross-Correlation, Effect-Size, Heterogeneity, Paired Design, Sampling Variance, Standard Deviation, Taylor's Law, Variance

Abstract

Meta-analyses are frequently used to quantify the difference in the average values of two groups (e.g., control and experimental treatment groups), but examine the difference in the variability (variance) of two groups. For such comparisons, the two relatively new effect size statistics, namely the log-transformed 'variability ratio' (the ratio of two standard deviations; lnVR) and the log-transformed 'CV ratio' (the ratio of two coefficients of variation; lnCVR) are useful. In practice, lnCVR may be of most use because a treatment may affect the mean and the variance simultaneously. We review current, and propose new, estimators for lnCVR and lnVR. We also present methods for use when the two groups are dependent (e.g., for cross-over and pre-test-post-test designs). A simulation study evaluated the performance of these estimators and we make recommendations about which estimators one should use to minimise bias. We also present two worked examples that illustrate the importance of accounting for the dependence of the two groups. We found that the degree to which dependence is accounted for in the sampling variance estimates can impact heterogeneity parameters such as τ^2 (i.e., the between-study variance) and I^2 (i.e., the proportion of the total variability due to between-study variance), and even the overall effect, and in turn qualitative interpretations. Meta-analytic comparison of the variability between two groups enables us to ask completely new questions and to gain fresh insights from existing datasets. We encourage researchers to take advantage of these convenient new effect size measures for the meta-analysis of variation.

1. INTRODUCTION

Meta-analysis is often used to evaluate studies comparing the average of two groups. These are usually treatment groups in an experiment/trial, one being a concurrent control, but may also represent naturally occurring groups (e.g., different sexes). The standardised mean difference (SMD; also known as Cohen's *d* and its associated derivatives), which is the difference between group means divided by the within-study variability, is a commonly-used effect size measure for this purpose [1]. SMD is popular because it is 'unitless', meaning it can be used to compare the results of studies that report outcomes in different units [2]. A similar unitless effect size measure, which can also be used to compare the means of two groups, is the logarithm of the ratio of the means of the groups. This effect size measure is known as the ratio of means in medicine (ROM [3]) and the log response ratio in ecology and evolution (lnRR [4]). Throughout we follow the lnRR notation as this will help to draw parallels with other effect size measures as we progress, but the reader should not be confused with the (logarithm of) risk ratio, which is also sometimes denoted (ln)RR. Surveys have shown that lnRR is the most widely used effect size measure in ecology and evolution [5-7]. Moreover, SMD and lnRR collectively account for over half of all meta-analyses in ecology [6,7], meaning comparisons between group means is the most widespread aim of meta-analysis in this field. SMD also seems to be among the most used standardised effect statistics in the medical and social sciences [8].

Two groups may not only differ in terms of their means, but also their variances [9]. At the most basic level, experimental treatments may directly increase or decrease the total amount of variance in a system due to inter-individual variability in response. Many biological systems also appear to display a mean-variance relationship [10-12]; most commonly, increasing averages are associated with increasing variances. Perhaps the most well-known example of a

biological mean-variance relationship comes from ecology and is known as Taylor's Law. This 'law' has been widely observed, and states that as mean population density increases, variance in population density also increases [13,14]. Where mean-variance relationships are present, a treatment may indirectly cause groups to have differing variances by altering the mean.

Nakagawa, Poulin, Mengersen, et al. [15] proposed a number of methods that allow the user to test for differences in the variance of groups meta-analytically. Among the methods proposed, the logarithm of the ratio of the standard deviations (SDs), named log 'variability ratio' (lnVR) and the logarithm of the ratio of the coefficients of variation (CV), termed the log 'CV ratio' (lnCVR), are most readily integrated into the standard meta-analytic paradigm; i.e. a contrast-based model using an effect size that corresponds to an effect relative to a concurrent control [16,17]. Of the two, lnCVR is perhaps the more useful measure where a mean-variance relationship is likely to exist. Nakagawa, Poulin, Mengersen, et al. [15] highlight that meta-analysing variation may be used to answer completely novel questions, but it can also be used to provide fresh insights into the topics on which a meta-analysis of means was already conducted. Indeed, lnCVR has already been applied in such diverse fields as ecology [18], evolution [19], agriculture [20], health [17], and the social sciences [21]. It is important to note that lnCVR (and also lnVR) require the same data to calculate as is already needed for computing SMD or lnRR values.

Our aims in this paper are threefold. First, we review existing and propose new estimators for lnCVR and its sampling error variance. These include, for the first time, estimators of the sampling variance when the two groups (treatment and control) are not independent (as may occur, for example, in cross-over trials or in paired, single-subject, or pre-test-post-test

96 designs). Second, we conduct a simulation study to investigate the performance of the
97 different estimators. Finally, we present two case studies using these techniques, and
98 illustrate the importance of accounting for dependence between the two treatment groups in
99 the estimation of sampling variation and other heterogeneity parameters (e.g., τ^2, the
100 between-study variance, and I^2 [22]).
101

102 ## 2. METHODS

103 ### 2.1 Point estimators when groups are independent

104 Let $x_T \sim N(\mu_T, \sigma_T)$ and $x_C \sim N(\mu_C, \sigma_C)$ denote normally distributed random variables with
105 true means (i.e., expected values) given by μ_T and μ_C and true standard deviations σ_T and σ_C.
106 For independent random samples based on these variables (e.g., representing some outcome
107 of interest measured in a treatment and control group) of size n_T and n_C, let $\bar{x}_T$ and $\bar{x}_C$ denote
108 the respective sample means and s_T and s_C the corresponding standard deviations for the two
109 groups. Then comparisons between the means, variances and coefficients of variation for two
110 groups can be made using the lnRR, lnVR and lnCVR effect size measures, respectively.
111 "Naïve" estimators of these effect statistics are:

$$\ln\mathrm{RR}_1 = \ln\left(\frac{\bar{x}_T}{\bar{x}_C}\right),$$

$$\ln\mathrm{VR}_1 = \ln\left(\frac{s_T}{s_C}\right),$$

$$\ln\mathrm{CVR}_1 = \ln\left(\frac{\mathrm{CV}_T}{\mathrm{CV}_C}\right),$$

112 where ln denotes the natural logarithm and $\mathrm{CV}_T = s_T/\bar{x}_T$ and $\mathrm{CV}_C = s_C/\bar{x}_C$ denote the
113 coefficients of variation in the treatment and control group, respectively.
114

While these naïve estimators are consistent and asymptotically unbiased, we can add corrections for the sample size based on a second-order Taylor expansion (also known as, the second order delta method) for each statistic [15,23,24]. For the lnRR, Lajeunesse [23] demonstrated such a correction is important to obtain unbiased estimation especially when sample size is small;

$$\mathrm{lnRR}_2 = \ln\left(\frac{\overline{x}_T}{\overline{x}_C}\right) + \frac{1}{2}\left(\frac{s_T^2}{n_T\overline{x}_T^2} - \frac{s_C^2}{n_C\overline{x}_C^2}\right).$$

Similarly, for the lnVR, Nakagawa, Poulin, Mengersen, et al. [15] proposed:

$$\mathrm{lnVR}_2 = \ln\left(\frac{s_T}{s_C}\right) + \frac{1}{2}\left(\frac{1}{n_T - 1} - \frac{1}{n_C - 1}\right).$$

Combing lnRR_2 and lnVR_2, one obtains:

$$\mathrm{lnCVR}_2 = \ln\left(\frac{\mathrm{CV}_T}{\mathrm{CV}_C}\right) + \frac{1}{2}\left(\frac{1}{n_T - 1} - \frac{1}{n_C - 1}\right) + \frac{1}{2}\left(\frac{s_C^2}{n_C\overline{x}_C^2} - \frac{s_T^2}{n_T\overline{x}_T^2}\right).$$

Note that Nakagawa, Poulin, Mengersen, et al. [15] originally suggested the an estimator of lnCVR that missed the bias correction pertaining to lnRR (i.e. $\frac{1}{2}\left(\frac{s_C^2}{n_C\overline{x}_C^2} - \frac{s_T^2}{n_T\overline{x}_T^2}\right)$). We also note that an alternative estimator of lnCVR could also be obtained based on $\left(1 + \frac{1}{4n}\right)\mathrm{CV}$, which it has been suggested acts as a 'rough' bias correction for the CV (e.g. [25]). However, this estimator is not recommended here, and it does not perform well (see Supplementary Materials S1, Text S1).

2.2 Dispersion estimators when the two groups are independent

The original estimators of the sampling (error) variance for lnRR [4] and lnVR [15] are based on the first-order Taylor expansion; they are respectively:

$$s^2(\mathrm{lnRR}_1) = \frac{s_C^2}{n_C\overline{x}_C^2} + \frac{s_T^2}{n_T\overline{x}_T^2},$$

$$s^2(\mathrm{lnVR}_1) = \frac{1}{2}\left(\frac{1}{n_C - 1} + \frac{1}{n_T - 1}\right).$$

Based on these, for lnCVR Nakagawa, Poulin, Mengersen, et al. [15] proposed:

$$s^2(\mathrm{lnCVR}_1) = \frac{s_C^2}{n_c\bar{x}_C^2} + \frac{1}{2(n_c - 1)} - 2\rho\sqrt{\frac{s_C^2}{n_c\bar{x}_C^2}\frac{1}{2(n_c - 1)}}$$

$$+ \frac{s_T^2}{n_T\bar{x}_T^2} + \frac{1}{2(n_T - 1)} - 2\rho\sqrt{\frac{s_T^2}{n_T\bar{x}_T^2}\frac{1}{2(n_T - 1)}},$$

where ρ is the correlation between the log mean and log SD. Nakagawa, Poulin, Mengersen, et al. [15] suggested that ρ can be estimated based on the correlation between the log sample mean and log sample SD across the studies included in a meta-analysis. However, in doing so one risks conflating within- and between-study correlation (i.e., the correlation in the bivariate sampling distribution of the sample mean and sample SD could be very different than the correlation of the true means and SDs across studies). In fact, for observations that come from an underlying population distribution that is symmetric (e.g. a normal distribution), the sample mean and variance are uncorrelated [26]. Thus, for the case considered here where $\rho = 0$ the equation above simplifies to:

$$s^2(\mathrm{lnCVR}_1) = \frac{s_C^2}{n_C\bar{x}_C^2} + \frac{1}{2(n_C - 1)} + \frac{s_T^2}{n_T\bar{x}_T^2} + \frac{1}{2(n_T - 1)}.$$

As a better estimator for the sampling variance of lnRR, Lajeunesse [23] derived and tested the following sampling variance based on the second-order Taylor expansion:

$$s^2(\mathrm{lnRR}_2) = \frac{s_C^2}{n_C\bar{x}_C^2} + \frac{s_C^4}{2n_C^2\bar{x}_C^4} + \frac{s_T^2}{n_T\bar{x}_T^2} + \frac{s_T^4}{2n_T^2\bar{x}_T^4}.$$

Similarly, we can derive the following sampling variance for lnVR based on the second-order Taylor expansion as:

$$s^2(\mathrm{lnVR}_2) = \frac{1}{2}\left(\frac{1}{n_C - 1} + \frac{1}{(n_C - 1)^2} + \frac{1}{n_T - 1} + \frac{1}{(n_T - 1)^2}\right).$$

Accordingly, the complete estimator of the sampling variance for lnCVR, based on $s^2(\text{lnRR}_2)$ and $s^2(\text{lnVR}_2)$ is:

$$s^2(\text{lnCVR}_2) = \frac{s_C^2}{n_C\bar{x}_C^2} + \frac{s_C^4}{2n_C^2\bar{x}_C^4} + \frac{1}{2(n_C-1)} + \frac{1}{2(n_C-1)^2}$$

$$+ \frac{s_T^2}{n_T\bar{x}_T^2} + \frac{s_T^4}{2n_T^2\bar{x}_T^4} + \frac{1}{2(n_T-1)} + \frac{1}{2(n_T-1)^2}.$$

In the supplementary materials, we propose estimators of the sampling covariance based on the above, which can be used when multiple treatment groups are contrasted with the same control [27] (see Supplementary Materials S1, Text S2)

2.3 Point estimators when groups are dependent

Due to experimental design, control and treatment groups are often not independent of one another. A clear example of this dependency is in the case of a cross-over design where the same individuals are subjected to both control and experimental treatments at two different time points. The point estimates given above will perform the same way regardless of whether we are dealing with independent or dependent groups. In cross-over studies, however, $n_T = n_C \equiv n$, unless dropouts are included in a pre-post design, in which case we recommend that $n = n_{post}$ (i.e. the sample size in the post-treatment condition) is used. This is because the correlation between pre and post-treatment measurements can only be calculated based on n, which assumes $n_T = n_C$ (see the next section). We can rewrite the dependent cases of lnRR_1 and lnRR_2 as:

$$\text{lnRR}_3 = \ln\left(\frac{\bar{x}_T}{\bar{x}_C}\right),$$

$$\text{lnRR}_4 = \ln\left(\frac{\bar{x}_T}{\bar{x}_C}\right) + \frac{1}{2}\left(\frac{s_T^2}{n\bar{x}_T^2} - \frac{s_C^2}{n\bar{x}_C^2}\right),$$

where subscripts 3 and 4 indicate the naïve estimator and estimator based the second-order Taylor expansion, respectively. Similarly, for lnVR and lnCVR, we have:

$$\ln\mathrm{VR}_3 = \ln\left(\frac{s_T}{s_C}\right),$$

$$\ln\mathrm{VR}_4 = \ln\left(\frac{s_T}{s_C}\right) + \frac{1}{2}\left(\frac{1}{n_T - 1} - \frac{1}{n_C - 1}\right) = \ln\left(\frac{s_T}{s_C}\right),$$

$$\ln\mathrm{CVR}_3 = \ln\left(\frac{\mathrm{CV}_T}{\mathrm{CV}_C}\right),$$

$$\ln\mathrm{CVR}_4 = \ln\left(\frac{\mathrm{CV}_T}{\mathrm{CV}_C}\right) + \frac{1}{2}\left(\frac{s_C^2}{n\bar{x}_C^2} - \frac{s_T^2}{n\bar{x}_T^2}\right).$$

2.4 Dispersion estimators when the two groups are dependent

In dependent cases estimates of the sampling variance need to account for the correlation between measurements from the same replicates on the two occasions (i.e. cross-correlation [28]). Based on the first-order Taylor expansion, the sampling variance for lnRR is:

$$s^2(\ln\mathrm{RR}_3) = \frac{s_C^2}{n_C\bar{x}_C^2} + \frac{s_T^2}{n_T\bar{x}_T^2} - 2r_{CT}\sqrt{\frac{s_C^2}{n_C\bar{x}_C^2}}\sqrt{\frac{s_T^2}{n_T\bar{x}_T^2}},$$

where r_{CT} is a cross-context correlation value estimated from the two sets of measurements on the same replicate when they are under the control and treatment conditions [29]. As discussed above for dependent studies $n_T = n_C \equiv n$ meaning $s^2(\ln\mathrm{RR}_3)$ simplifies to:

$$s^2(\ln\mathrm{RR}_3) = \frac{s_C^2}{n\bar{x}_C^2} + \frac{s_T^2}{n\bar{x}_T^2} - r_{CT}\frac{2s_Cs_T}{n\bar{x}_C\bar{x}_T}.$$

If based on the second-order Taylor expansion [23], the estimator of the sampling variance for lnRR is:

$$s^2(\ln\mathrm{RR}_4) = \frac{s_C^2}{n\bar{x}_C^2} + \frac{s_T^2}{n\bar{x}_T^2} - r_{CT}\frac{2s_Cs_T}{n\bar{x}_C\bar{x}_T} + \frac{s_C^4}{2n^2\bar{x}_C^4} + \frac{s_T^4}{2n^2\bar{x}_T^4} + r_{CT}^2\frac{s_C^2s_T^2(\bar{x}_C^4 + \bar{x}_T^4)}{2n^2\bar{x}_C^4\bar{x}_T^4}.$$

We can also derive the sampling variance for dependent cases of lnVR based on the first-order Taylor expansion as:

$$s^2(\ln \mathrm{VR}_3) = \frac{1}{2}\left(\frac{1}{(n_C-1)} + \frac{1}{(n_T-1)}\right) - r_{CT}^2\sqrt{\frac{1}{(n_C-1)}}\sqrt{\frac{1}{(n_T-1)}},$$

176 which, where $n_T = n_C \equiv n$, simplifies to:

$$s^2(\ln \mathrm{VR}_3) = \frac{1-r_{CT}^2}{n-1}.$$

177 Based on the second-order Taylor expansion, we have the sampling variance for dependent

178 cases of lnVR as:

$$s^2(\ln \mathrm{VR}_4) = \frac{1}{n-1} - r_{CT}^2\frac{1}{n-1} + \frac{1}{(n-1)^2} + r_{CT}^4\frac{s_C^8+s_T^8}{2(n-1)^2 s_C^4 s_T^4}.$$

179 From the sampling variances for lnRR and lnVR, we have the sampling variance for lnCVR

180 with first- and second-order Taylor expansion as:

$$s^2(\ln \mathrm{CVR}_3) = \frac{s_C^2}{n\bar{x}_C^2} + \frac{s_T^2}{n\bar{x}_T^2} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T} + \frac{1}{n-1} - r_{CT}^2\frac{1}{n-1},$$

$$s^2(\ln \mathrm{CVR}_4) = \frac{s_C^2}{n\bar{x}_C^2} + \frac{s_T^2}{n\bar{x}_T^2} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T} + \frac{s_C^4}{2n^2\bar{x}_C^4} + \frac{s_T^4}{2n^2\bar{x}_T^4} + r_{CT}^2\frac{s_C^2 s_T^2(\bar{x}_C^4+\bar{x}_T^4)}{2n^2\bar{x}_C^4\bar{x}_T^4}$$

$$+\frac{1}{n-1} - r_{CT}^2\frac{1}{n-1} + \frac{1}{(n-1)^2} + r_{CT}^4\frac{s_C^8+s_T^8}{2(n-1)^2 s_C^4 s_T^4}.$$

181 Note that, where *r* is positive the estimated sample variance for a dependent estimator will be

182 smaller than its independent equivalent, but that as *r* shrinks to 0 the dependent case

183 converges on the independent; e.g. assuming $n_c = n_T$, where $r > 0$, $s^2(\ln \mathrm{CVR}_3) < s^2(\ln \mathrm{CVR}_1)$,

184 but where $r = 0$, $s^2(\ln \mathrm{CVR}_3) = s^2(\ln \mathrm{CVR}_1)$.

185

186 ## 3. SIMULATION

187 ### 3.1 Simulation study design

188 We simulated a two-group experiment/trial, where a pair of groups is based on n_T and n_C

189 random samples drawn from populations under an experimental treatment and control

190 conditions. The treatment and control populations have means μ_T and μ_C and standard

191 deviations (SDs) σ_T and σ_C, respectively. The *i*th sample in each group, y_{Ti} ($i = 1 \ldots n_T$) and

192 y_{Ci} ($i = 1 \ldots n_C$) was drawn from a bivariate normal distribution as follows:

$$\begin{pmatrix} y_{Ti} \\ y_{Ci} \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_T \\ \mu_C \end{bmatrix}, \begin{bmatrix} \sigma_T^2 & \rho_{CT}\sigma_T\sigma_C \\ \rho_{CT}\sigma_C\sigma_T & \sigma_C^2 \end{bmatrix} \right)$$

193 Where $\begin{bmatrix} \mu_T \\ \mu_C \end{bmatrix}$ are the population means of the two groups, $\begin{bmatrix} \sigma_T^2 & \rho_{CT}\sigma_T\sigma_C \\ \rho_{CT}\sigma_C\sigma_T & \sigma_C^2 \end{bmatrix}$ is a variance

194 co-variance matrix specifying the variances of the two groups with ρ_{CT} giving the degree of

195 correlation among the *i*th samples in the two groups and all other parameters are as above.

196 When where $\rho_{CT} \neq 0$ the *i*th data in the two groups are correlated (i.e. dependent or paired

197 samples as in a cross-over design).

198

199 In all simulations, $\mu_C = 100$ and $\sigma_C = 20$, which across the parameters tested ensures positive

200 sample means (required for log transformation). We explored values of μ_T ranging between

201 $\mu_C \times e^{-0.5}$ and $\mu_C \times e^{0.5}$ and values of σ_T ranging between $\sigma_C \times e^{-0.5}$ and $\sigma_C \times e^{0.5}$, meaning the

202 $\ln(\mu_T / \mu_C)$ and $\ln(\sigma_T / \sigma_C)$ is between -0.5 and 0.5. All combinations were explored and

203 where $\ln(\mu_T / \mu_C) = \ln(\sigma_T / \sigma_C)$ the coefficient of variance (CV) of the two groups will be

204 identical. We explored $n_C = 8$, 16 and 42, with $n_C = n_T$ and, with $n_C < n_T$ (independent case).

205 We also explored $\rho_{CT} = 0$ and $\rho_{CT} = 0.8$. For each set of parameters, we simulated 100,000

206 experiments.

207

208 Based on the sample means and SDs of each simulated experiment, we calculated $\ln CVR_1$

209 and $\ln CVR_2$ for independent cases ($\rho_{CT} \neq 0$) and $\ln CVR_3$ and $\ln CVR_4$ for dependent cases

210 ($\rho_{CT} \neq 0$). We also calculated the sampling variance estimators $s^2(\ln CVR_1)$ and $s^2(\ln CVR_2)$

211 where $\rho_{CT} \neq 0$, and $s^2(\ln CVR_3)$ and $s^2(\ln CVR_4)$ where $\rho_{CT} \neq 0$. We calculated bias in the *i*th

212 estimator as:

$$\text{bias}[\ln\text{CVR}_i] = \frac{1}{K}\sum_{k=1}^{K} \ln\text{CVR}_{ik} - \ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right),$$

where *k* is the *k*th value of *K* (here 100,000) simulated values of $\ln\text{CVR}_i$ ($k = 1 \ldots K$). This bias can be interpreted as the mean deviation of the *i*th estimator of lnCVR from the true population value. We calculated bias in sampling variance estimator *i* as:

$$\text{bias}[s^2(\ln\text{CVR}_i)] = \frac{s^2(\ln\text{CVR}_i) - \theta_j^2}{\theta_j^2} \times 100,$$

where $s^2(\ln\text{CVR}_i)$ is the value of the *i*th sampling variance based on the simulated population statistics and sample sizes and θ_j is the SD among *K* simulated effect sizes estimated using estimator *j*. This bias can be interpreted as the percentage by which the sampling variance estimator deviates from the true value (i.e. 100 = the estimator is twice the true value). We calculated coverage as the proportion of 95% confidence intervals (CIs) that include $\ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right)$. For a combination of the *j*th effect size estimator ($\ln\text{CVR}_j$) and *i*th sampling variance $s^2(\ln\text{CVR}_i)$, 95% CIs were constructed as:

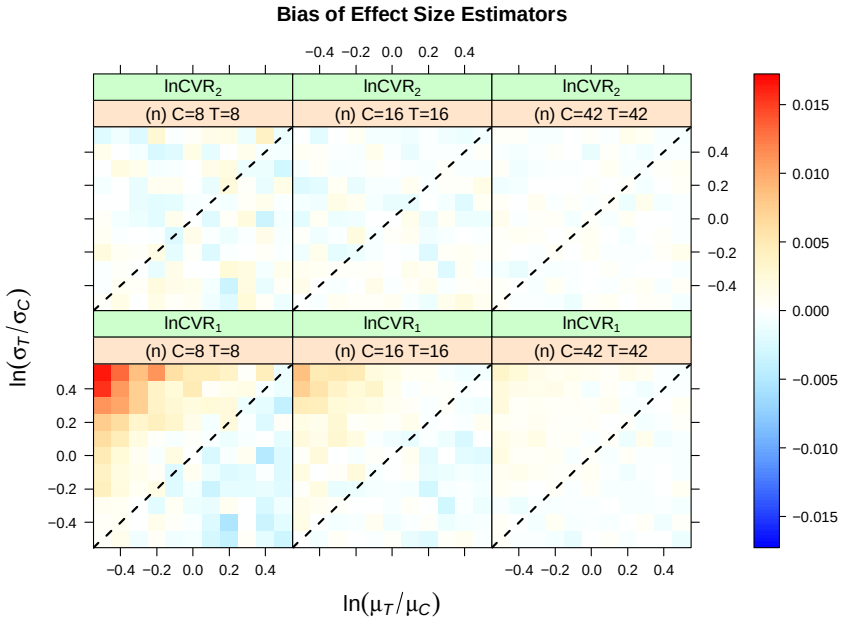
$$95\%\ \text{CI} = \ln\text{CVR}_j \pm z_{0.975}s(\ln\text{CVR}_i)$$

where $\ln\text{CVR}_j$ is the estimated effect size for the simulated sample, $s(\ln\text{CVR}_i)$ an estimate of the standard error (SE; the square root of the estimated sampling variance), and $z_{0.975}$ is the function of the 0.975[th] quantile of a *z* distribution (approx. 1.96). Simulations and analyses were performed in R v3.5.1; [30], and using the 'mvrnorm' function in the *MASS* package [31]. All data and code presented in this manuscript can be found at (https://github.com/AlistairMcNairSenior/lnCVR_Estimators_Sim).

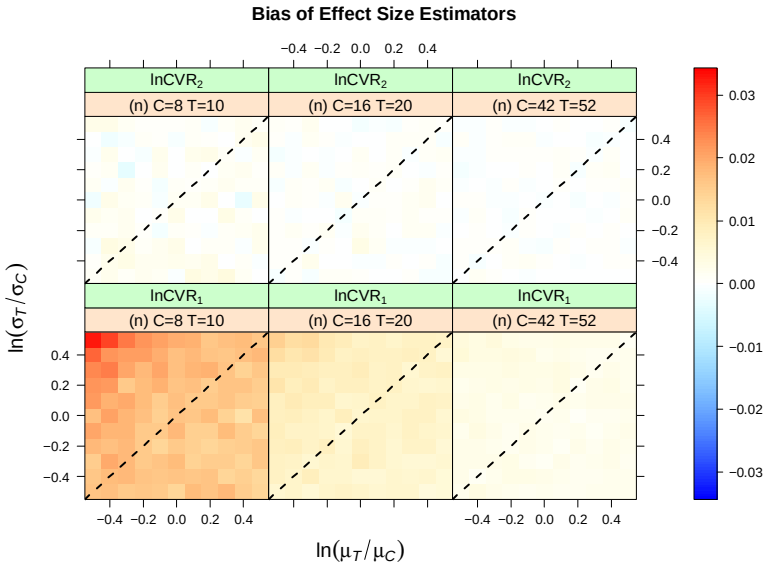
3.2 Simulation results

We begin with the case where the two groups are independent ($\rho_{CT} = 0$). Figure 1 shows bias in the estimated effect as a function of sample size and the log the ratio of the means and SDs

in the two groups. Across the diagonal elements of each plot (black-dashed line) the underlying CV of the two populations is identical (even if the means and SDs differ; $\ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right) = 0$), elements above the line correspond to the CV of the treatment population being greater than that of the control group ($\ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right) > 0$), and elements below the line the opposite ($\ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right) < 0$). $\ln CVR_1$ overestimates positive effects and slightly under-estimate negative effects, with bias being most profound where the sample size is small. $\ln CVR_2$, on the other hand, displays no systematic bias. Figure 2 shows the results where the sample size of the treatment group is ~25% greater than that of the control group. $\ln CVR_1$ showed severe upward bias, especially where the sample size was small, where as $\ln CVR_2$ performed with only very minor upward bias, which all but disappeared for larger sample sizes. Given that $\ln CVR_2$ was determined to be the most accurate estimator of the effect, we proceeded to explore how $\ln CVR_2$ performed in conjunction with different estimators of sampling variance.



248 **Figure 1:** Bias in effect size estimators of lnCVR as a function of the log ratio of population
249 means (*x*-axis), SDs (*y*-axis) and sample size (balanced) for the case of independent treatment
250 and control group data (ρ_{CT} = 0). Black dashed line indicates no effect (i.e., lnCVR = 0).
251



252

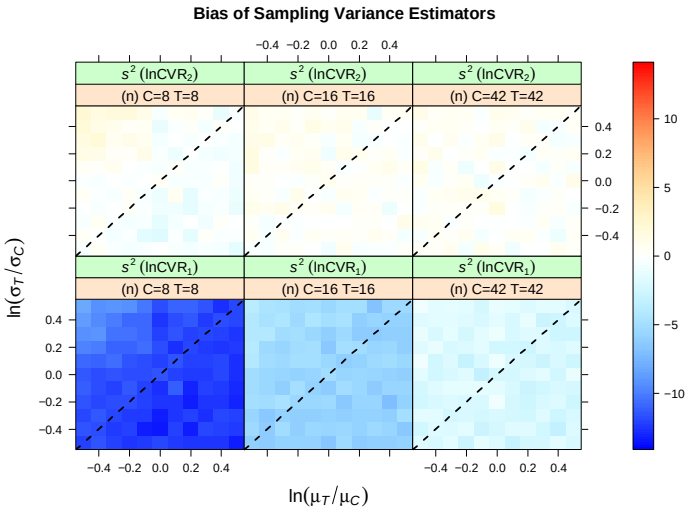
253 **Figure 2:** Bias in effect size estimators of lnCVR as a function of the log ratio of population
254 means (*x*-axis), SDs (*y*-axis) and sample size (unbalanced) for the case of independent
255 treatment and control group data (ρ_{CT} = 0). Black dashed line indicates no effect (i.e., lnCVR
256 = 0).
257

258 The first sampling variance estimator $s^2(\text{lnCVR}_1)$ underestimated the variance among
259 simulated values of lnCVR_2, particularly where the sample size was small (Figure 3). Biases
260 for $s^2(\text{lnCVR}_2)$ were minimal, although there was some very slight upward bias for small
261 sample sizes and large positive effects (Figure 3). The coverage of 95% CIs for $s^2(\text{lnCVR}_1)$
262 and $s^2(\text{lnCVR}_2)$ (paired with lnCVR_2) are shown in Figure 4. $s^2(\text{lnCVR}_1)$ generated CIs that
263 were too narrow at smaller sample sizes, whereas again $s^2(\text{lnCVR}_2)$ performed with little
264 bias. At larger sample sizes coverage was much closer to the nominal level (Figure 4),

265 although $s^2(\text{lnCVR}_2)$ still performed more accurately. The same patterns of performance were

266 observed for the case where $n_C < n_T$ (Supplementary Figures S1 and S2).

267



268

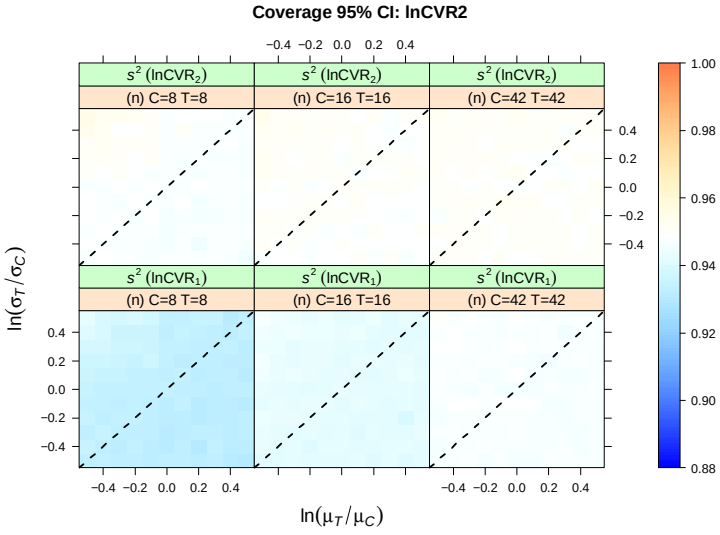
269 **Figure 3:** Bias in sampling variance estimators of lnCVR as a function of the log ratio of

270 population means (*x*-axis), SDs (*y*-axis) and sample size (balanced) for the case of

271 independent treatment and control group data (ρ_{CT} = 0). Black dashed line indicates no effect

272 (i.e., lnCVR = 0).

273



274

Figure 4: Coverage of 95% CIs based on estimators of the sampling variance of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis) and sample size (balanced) for the case of independent treatment and control group data (ρ_{CT} = 0). Black dashed line indicates no effect (i.e., lnCVR = 0).

For the case where treatment and control samples were dependent on one another (ρ_{CT} = 0.8) $lnCVR_4$ out-performed $lnCVR_3$, with a pattern identical to that in Figure 1 (Figure S3). With regards the two estimators for dependent sampling variances, $s^2(lnCVR_3)$ underestimated the variance where as $s^2(lnCVR_4)$ overestimated the variance (Figure 5). These biases were within a reasonable range for larger samples, but were severe for small samples, and $s^2(lnCVR_4)$ in particular showed extreme upward bias (reaching 60% overestimate) when the SD of the treatment group differed from that of the control group (Figure 5). The CIs generated by $s^2(lnCVR_3)$ had a tendency to be too narrow whereas those generated by $s^2(lnCVR_4)$ were too wide (Figure 6).

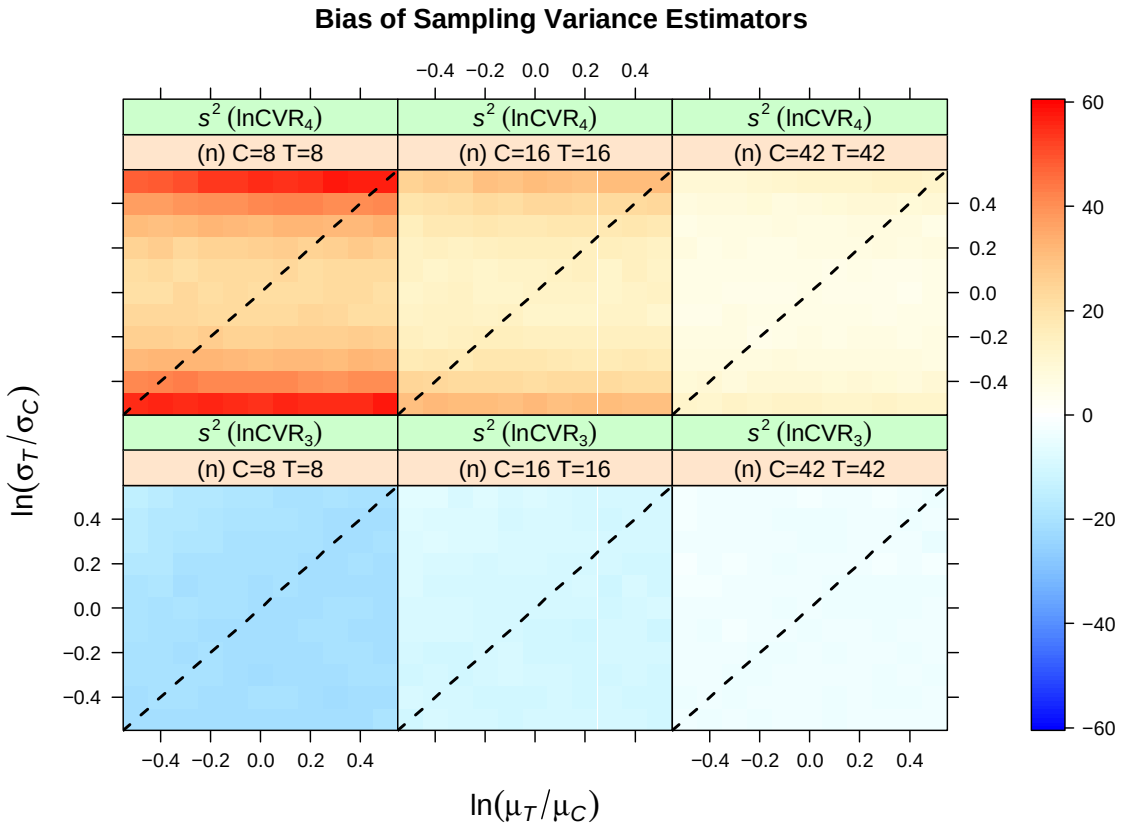


Figure 5: Bias in sampling variance estimators of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis) and sample size (balanced) for the case of dependent treatment and control group data (ρ_{CT} = 0.8). Black dashed line indicates no effect (i.e., lnCVR = 0).

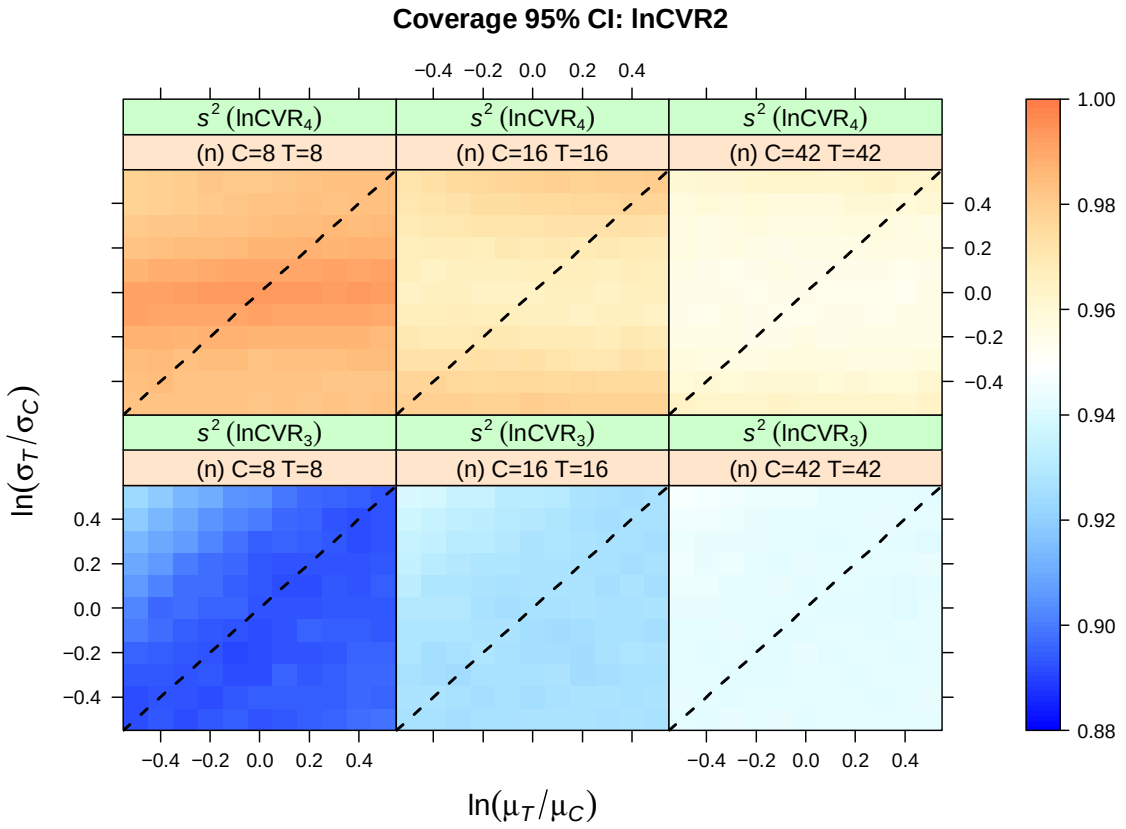


Figure 6: Coverage of 95% CIs based on estimators of the sampling variance of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis) and sample size (balanced) for the case of dependent treatment and control group data (ρ_{CT} = 0.8). Black dashed line indicates no effect (i.e., lnCVR = 0).

4. WORKED EXAMPLES

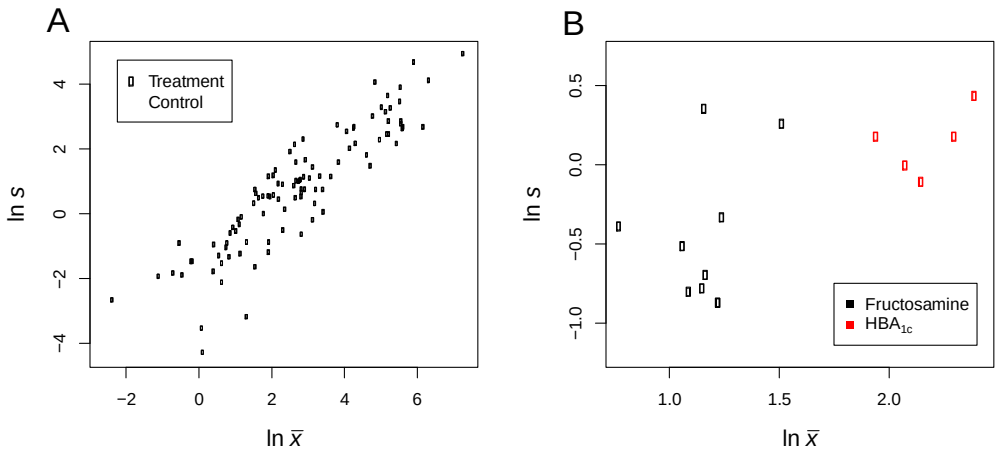
We now provide two examples: one from the field of ecology and the other from the health sciences. All meta-analytic models (random-effects meta-analysis) were fitted using the 'rma' function (with default settings) in *metafor* [32].

4.1 Example 1: Carbon dioxide levels and plant mass

Curtis, Wang [33] performed a meta-analysis of experimental studies that tested for the effects of elevated carbon dioxide (CO_2) levels on woody plant mass. Briefly, these studies

compared the total biomass (above and below ground) of plants grown under ambient and artificially elevated (~100% increase) CO_2 levels. Studies were performed in a range of contexts, including highly controlled (e.g., green houses) and less controlled (e.g., field sites) environments, as well as across temperature, light, water, and soil-fertility levels. Replication was at the level of the locale (e.g., plot/site/greenhouse) at which a treatment was applied, and treatment/control groups may be correlated (i.e., non-independent) if, for example, locales experiencing different treatments are paired spatially or temporally. However, the degree to which such correlations are present was not stated. Aggregating 102 effect sizes (lnRR), Curtis, Wang [33] found that the mean biomass of woody plants at a site increases by, on average, 28.8% under elevated CO_2 conditions. However, there was evidence that the effect is moderated by the presence of other stressors such as under nutrient- or light-limited conditions.

Here we ask whether elevated CO_2 levels also increase among-replicate variability in plant biomass using lnCVR. We tested the sensitivity of the analysis to the assumption that treatment and control groups are uncorrelated. Because we do not know precisely which effect size data come from paired designs, we calculated effect sizes and sampling variance assuming complete independence (0% of effect sizes have correlated groups), varying degrees of partial dependence (a random subset of 20%, 60%, or 80% effect sizes have correlated groups; $r_{CT} = 0.8$), or complete dependence (100% of effect sizes have correlated groups; $r_{CT} = 0.8$). For those effect sizes that were assumed to be uncorrelated we used lnCVR_2 and $s^2(\text{lnCVR}_2)$, and for those that are correlated lnCVR_4 and $s^2(\text{lnCVR}_3)$.



333

334 **Figure 7:** Association between log sample mean (ln $\bar{x}$) and log sample standard deviation (ln
335 *s*) for treatment (hollow points) and control (solid points) groups in the data from; A) Curtis,
336 Wang [33], where the outcome is woody plant biomass under elevated (treatment) *vs* ambient
337 (control) CO_2 levels; and B) Brand-Miller et al. (2003) where the outcome is a measure of
338 glycemia in diabetic individuals on low (treatment) *vs* high (control) glycemic index diets.
339 Note in (B) measures of glycemia are either fructosamine (black points) or HbA_{1c} (red points)
340 levels, where lower levels indicate better gylcemic control.

341

342 There was evidence for a mean-variance relationship under both elevated and ambient CO_2
343 levels (Figure 7A). The influence of increasing the percentage of effect sizes that are assumed
344 to come from correlated groups on a random-effects meta-analysis is shown in Table 1. There
345 are some qualitative differences in the interpretation of the overall effect, whereby the
346 associated CI spans zero in some cases, but not others (Table 1). In all cases the sign of the
347 overall effect is stable and suggests that elevating CO_2 levels on average decreases the CV in
348 biomass among replicates (possibly by somewhere between $100 \times (1 - \exp(-0.078) = 7.5$ to
349 $100 \times (1 - \exp(-0.116) = 10.9$ percent). The effect of increasing the number of studies with
350 correlated groups on the estimated inter-effect size heterogeneity, is however, much more

dramatic. As less independence is assumed, the amount of heterogeneity (absolutely, in terms of $\hat{\tau}^2$, and relatively, in terms of I^2) increases substantially (Table 1), such that when 40% or more of the studies are have used paired designs, Cochran's *Q* test yields a significant result.

4.2 Example 2: Low glycemic index diets and glycemic control in diabetic subjects

Brand-Miller, Petocz, Hayne, Colagiuri [34] performed a meta-analysis of studies designed to test the effects of low glycemic index (GI) diets on bio-markers of glycemic control in diabetic (type 1 and 2) individuals. Individuals were given either low or high GI diets, after which glycemia was measured using HbA_{1c} and/or fructosamine levels. These two markers quantify glycemia over longer *vs* shorter time periods respectively, where lower levels indicate better glycemic control. The studies differed somewhat in the overall GI of the diets used and the duration for which subjects were on the diets. The studies used a mixture of parallel designs where the individuals in each treatment group are completely independent, and cross-over designs where each individual was subject to both treatments. Brand-Miller, Petocz, Hayne, Colagiuri [34] acknowledged that for those studies with a cross-over design, there will be a degree of correlation among the treatment and control condition data. They tested the sensitivity of their results to any such correlation by repeating the analyses assuming complete independence ($r_{CT} = 0$) and also assuming that groups are correlated ($r_{CT} = 0.34$; based on one of the studies in their primary literature). Their analyses of 14 effect sizes (mean differences, expressed in terms of percent; 11 from studies with cross-over designs) suggested that measures of glycemia are decreased by 6.8 percentage points (improved glycemic control) on low GI diets irrespective of their assumptions about correlations among groups. The authors used a fixed-effect meta-analytic model, and did not present heterogeneity statistics.

We tested whether low GI diets affect inter-individual variability in glycemic control using lnCVR. Unlike example 1, here we do know which studies contain dependent groups (those with cross-over designs), although the strength of the dependence is not precisely known. For independent designs we calculated effect sizes and sampling variances *via* $\ln CVR_2$, and $s^2(\ln CVR_2)$. For those studies using a cross-over design we calculated $\ln CVR_4$ and $s^2(\ln CVR_3)$ assuming treatment and control data are correlated with r_{CT} = 0, 0.3, 0.5, and 0.8. Where more than one measure of glycemia was presented from a single study, we primarily use fructosamine levels (this being the more widely reported measure).

We observed a mean-variance relationship amongst both measures of glycemic control within the two treatment groups (Figure 7B). The results of random-effects meta-analyses fitted to the effect sizes are given in Table 2. The analyses estimated that on low-GI diets the CV in biomarkers of glycemic control is on average reduced by between 13% ($100 \times (1 - \exp(-0.135))$) and 18% ($100 \times (1 - \exp(-0.177))$) compared to high-GI diets. However, as the degree of correlation among data from cross-over trials increased, there was a marginal reduction in the overall effect magnitude and an increase in the associated SE (Table 2); for r_{CT} = 0.5, the overall effect was not statistically significant. With increasing correlation, heterogeneity also increased (Table 2). Where we assumed complete independence (r_{CT} = 0), there was no evidence for heterogeneity, but for r_{CT} = 0.8, we detected inter effect size heterogeneity (Table 2).

5. DISCUSSION AND CONCLUSIONS

We recommend that meta-analysts use the following estimator of the lnCVR for independent study designs:

$$\ln\mathrm{CVR}_{\mathrm{ind}} = \ln\left(\frac{\mathrm{CV}_T}{\mathrm{CV}_C}\right) + \frac{1}{2}\left(\frac{1}{n_T - 1} - \frac{1}{n_C - 1}\right) + \frac{1}{2}\left(\frac{s_C^2}{n_C\bar{x}_C^2} - \frac{s_T^2}{n_T\bar{x}_T^2}\right).$$

For dependent study designs we recommend the use of the following point estimator:

$$\ln\mathrm{CVR}_{\mathrm{dep}} = \ln\left(\frac{\mathrm{CV}_T}{\mathrm{CV}_C}\right) + \frac{1}{2}\left(\frac{s_C^2}{n\bar{x}_C^2} - \frac{s_T^2}{n\bar{x}_T^2}\right).$$

Under the simulated conditions explored, these estimators exhibited minimal bias, where 'naïve' estimators displayed systematic biases, substantially overestimating large positive effects, especially when sample sizes were small. Compared to previous estimators [15], this revision contains an additional term, $\frac{1}{2}\left(\frac{s_C^2}{n_C\bar{x}_C^2} - \frac{s_T^2}{n_T\bar{x}_T^2}\right)$, which has also been shown to reduce bias in mean effects estimated *via* lnRR [23]. We also recommend that the following estimators for the sampling variance of lnCVR be used for independent and dependent study designs, respectively:

$$s^2(\ln\mathrm{CVR}_{\mathrm{ind}}) = \frac{s_C^2}{n_C\bar{x}_C^2} + \frac{s_C^4}{2n_C^2\bar{x}_C^4} + \frac{1}{2(n_C - 1)} + \frac{1}{2(n_C - 1)^2}$$

$$+ \frac{s_T^2}{n_T\bar{x}_T^2} + \frac{s_T^4}{2n_T^2\bar{x}_T^4} + \frac{1}{2(n_T - 1)} + \frac{1}{2(n_T - 1)^2},$$

$$s^2\left(\ln\mathrm{CVR}_{\mathrm{dep}}\right) = \frac{s_C^2}{n\bar{x}_C^2} + \frac{s_T^2}{n\bar{x}_T^2} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T} + \frac{1}{n-1} - r_{CT}^2\frac{1}{n-1}.$$

Our simulations demonstrate that the estimator for independent designs performs very well and 95% CIs based on a *z* distribution give coverage at the nominal level. The estimator for dependent cases slightly underestimates the actual sampling variance in lnCVR, and will generate CIs (based on *z* or *t* distributions) that are slightly too narrow. This might be due to the substitution of r_{CT} for the unknown true correlation in the equation for the sampling variance without further account of the additional source of uncertainty this introduces. CIs that are too narrow may be more troublesome in that they can lead to inflated type-1 error rates (a more conservative estimator, $s^2(\ln\mathrm{CVR}_4)$, is given above, although this approach may

substantially overestimate the sampling variance for small samples). Note that these recommended estimators are now available in the 'escalc' function in the development version of *metafor* (https://github.com/wviechtb/metafor), and will eventually be implemented in the CRAN version.

We used the recommended estimators to evaluate whether: 1) increased CO_2 levels affect variation in woody plant biomass, and 2) low-GI diets alter between-individual variation in glycemic control in diabetics. In both cases, we found that the treatments have a tendency to decrease the CV. In both cases the analyses were sensitive to assumptions about the degree to which treatment and control data are correlated. Assuming higher degrees of correlation resulted in small changes in the overall effect (and its standard error). Although these parameters were relatively stable, for estimates with CIs close to zero, changing assumptions about group independence can affect inference. Increasing the degree of correlation dramatically increased the estimated between-effect size heterogeneity, which could change conclusions about the consistency of the reported effects. This trend can be explained by the fact that as more/stronger correlations are assumed the sampling variances associated with the individual effect sizes shrink, effects are assumed to be more precise, and sampling variability therefore becomes less able to explain the variation among the effects. Our results corroborate the points made by Becker [28], who introduced an estimator for the sampling variance of SMD for dependent groups.

As is the case with any exercise in data analysis, the most appropriate technique to use will depend on the question being asked. Where the analyst is able to determine with a reasonable degree of certainty that a mean-variance relationship does not exist, lnVR may be a more useful measure of between-group differences in variability than lnCVR. This is because

lnCVR risks conflating effects on the SD with effects on the mean. In other instances, the user may be more interested in ascertaining whether a treatment alters the SD irrespective of a mean-variance relationship (e.g., in questions related to power and study design), and again lnVR would be an appropriate choice. However, where mean-variance relationships are present, and the analyst is interested in whether the variation is greater/lower than expected given the mean, lnCVR is useful. For some matters, it may even be common practice for the primary literature to describe variation in terms of CV rather than SD. For instance, in ecology and evolution it is common to present CV when comparing variability amongst species/traits that exist on different scales because CV is a relative measure [35]. We note that such a practice is not necessarily required for meta-analysis because lnVR is also a relative measure of variation, and as such should also do a good job of correcting for inter-system differences in scale. Nevertheless, where CV is the measure of variability commonly reported in the primary literature, the user may find it intuitive (or even necessary) to use lnCVR.

Nakagawa, Poulin, Mengersen, et al. [15] also present alternative arm-based models (and discuss bivariate models) for meta-analysis of variation. The lnCVR metric assumes that changes in the mean are associated with proportional changes in the SD. Arm-based (and bivariate) models are an alternative for meta-analysis which allow the user to circumvent the assumption of proportionality. Arm-based models, however, are not without their critics who argue that these methods are radical departure from established meta-analytic thinking (see [16]). Like other (contrast-based) effect size measures that reflect the difference between two groups (e.g., the standardized mean difference, log response ratio, log risk/odds ratio or the risk difference), lnCVR readily integrates with our most widespread analytical paradigms, offering a convenient and intuitive method for meta-analysis of variability.

Finally, we finish by reiterating the point made by Nakagawa, Poulin, Mengersen, et al. [15], and echoed by subsequent papers using lnCVR in different fields of study [17-21]. Meta-analysis of variation can tackle entirely new questions and open our eyes to insights that are hidden in datasets. The datasets required to gain these insights already exist because lnCVR is based on the same summary statistics as SMD and lnRR; means, SDs, and sample sizes. We suspect over 50,000 datasets of this sort have already been collected (c.f. [36]). In this regard it is vital that meta-analytic 'raw' data are made available and reusable in the spirit of open and transparent science [37,38].

ACKNOWLEDGEMENTS

AMS was supported by an Australian Research Council Discovery Early Career Researcher Award (ARC DECRA: DE180101520). SN was supported by an ARC Discovery Grant (DP180100818). We would like to thank Professor Jennie Brand-Miller for her thoughts on our analysis of glycemia in diabetic individuals.

REFERENCES

1. Hedges LV, Olkin I. *Statistical methods for meta-analysis.* San Diego: Academic Press; 1985.
2. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biol Rev.* 2007; 82(4): 591-605.
3. Friedrich JO, Adhikari NK, Beyene J. The ratio of means method as an alternative to mean differences for analyzing continuous outcome variables in meta-analysis: A simulation study. *BMC Med Res Methodol.* 2008; 8(32).
4. Hedges LV, Gurevitch J, Curtis PS. The meta-analysis of response ratios in experimental ecology. *Ecology.* 1999; 80(4): 1150-1156.
5. Senior AM, Grueber CE, Kamiya T, et al. Heterogeneity in ecological and evolutionary meta-analyses: Its magnitude and implications. *Ecology.* 2016; 97(12): 3293-3299.
6. Nakagawa S, Santos ESA. Methodological issues and advances in biological meta-analysis. *Evol Ecol.* 2012; 26(5): 1253-1274.
7. Koricheva J, Gurevitch J. Uses and misuses of meta-analysis in plant ecology. *J Ecol.* 2014; 102(4): 828-844.
8. Cooper H, Hedges LV, Valentine JC. *The handbook of research synthesis and meta-analysis.* Russell Sage Foundation; 2009.
9. Osenberg CW, Sarnelle O, Cooper SD. Effect size in ecological experiments: The application of biological models in meta□analysis. *Am Nat.* 1997; 150(6): 798-812.
10. Cohen JE, Xu M. Random sampling of skewed distributions implies taylor's power law of fluctuation scaling. *Proc Natl Acad Sci USA.* 2015; 112(25): 7749-7754.
11. Cohen JE, Xu M, Schuster WSF. Allometric scaling of population variance with mean body size is predicted from taylor's law and density-mass allometry. *Proc Natl Acad Sci USA.* 2012; 109(39): 15829-15834.
12. Nakagawa S, Schielzeth H. The mean strikes back: Mean-variance relationships and heteroscedasticity. *Trends Ecol Evol.* 2012; 27(9): 474-475.
13. Taylor LR. Aggregation, variance and the mean. *Nature.* 1961; 189(4766): 732-735.
14. Bartlett MS. Some notes on insecticide tests in the laboratory and in the field. *Supplement to the Journal of the Royal Statistical Society.* 1936; 3(2): 185-194.
15. Nakagawa S, Poulin R, Mengersen K, et al. Meta-analysis of variation: Ecological and evolutionary applications and beyond. *Methods Ecol Evol.* 2015; 6(2): 143-152.
16. Dias S, Ades AE. Absolute or relative effects? Arm-based synthesis of trial data. *Res Syn Methods.* 2016; 7: 23-28.
17. Senior AM, Gosby AK, Lu J, Simpson SJ, Raubenheimer D. Meta-analysis of variance: An illustration comparing the effects of two dietary interventions on variability in weight. *Evol Med Public Health.* 2016; 2016(1): 244-255.
18. Senior AM, Nakagawa S, Lihoreau M, Simpson SJ, Raubenheimer D. An overlooked consequence of dietary mixing: A varied diet reduces inter-individual variance in fitness. *Am Nat.* 2015; 186(5): 649-659.
19. Janicke T, Häderer IK, Lajeunesse MJ, Anthes N. Darwinian sex roles confirmed across the animal kingdom. *Science Advances.* 2016; 2(2).
20. Knapp S, van der Heijden MGA. A global meta-analysis of yield stability in organic and conservation agriculture. *Nat Commun.* 2018; 9(1): 3632.
21. O'Dea RE, Lagisz M, Jennions MD, Nakagawa S. Gender differences in individual variation in academic grades fail to fit expected patterns for stem. *Nat Commun.* 2018; 9(1): 3777.

22. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002; 21(11): 1539-1558.
23. Lajeunesse MJ. Bias and correction for the log response ratio in ecological meta-analysis. *Ecology.* 2015; 96(8): 2056-2063.
24. Raudenbush SW, Bryk AS. Examining correlates of diversity. *Journal of Educational Statistics.* 1987; 12: 241-269.
25. Sokal R, Rohlf FJ. *Biometry*. 3 ed. New York: Freeman; 1995.
26. Zhang L. Sample mean and sample variance. *American Statistician.* 2007; 6(2): 159-160.
27. Noble DWA, Lagisz M, O'Dea RE, Nakagawa S. Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Mol Ecol.* 2017; 26(9): 2410-2425.
28. Becker BJ. Synthesizing standardized mean□change measures. *Br J Math Stat Psychol.* 1988; 41(2): 257-278.
29. Lajeunesse MJ. On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology.* 2011; 92(11): 2049-2055.
30. *R: A language and environemnt for statistical computing* [computer program]. Version 3.5.1. Available at http://www.r-project.org.2018.
31. Venables WN, Ripley BD. *Modern applied statistics with s.* 4 ed. New York: Springer; 2002.
32. Viechtbauer W. Conducting meta-analyses in r with the metafor package. *J Stat Softw.* 2010; 36: 1-48.
33. Curtis PS, Wang X. A meta-analysis of elevated co2 effects on woody plant mass, form, and physiology. *Oecologia.* 1998; 113(3): 299-313.
34. Brand-Miller J, Petocz P, Hayne S, Colagiuri S. Low-glycemic index diets in the management of diabetes. *Diabetes Care.* 2003; 26(8): 2261-2267.
35. Van Valen L. The statistics of variation. *Evol Theory.* 1978; 4(433-443).
36. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature.* 2018; 8(555): 175-182.
37. Nosek BA, Alter G, Banks GC, et al. Promoting an open research culture. *Science.* 2015; 348(6242): 1422.
38. Parker TH, Forstmeier W, Koricheva J, et al. Transparency in ecology and evolution: Real problems, real solutions. *Trends Ecol Evol.* 2016; 31(9): 711-719.

564 **Table 1:** Estimates of the overall effect (lnCVR) and heterogeneity statistics from random-
565 effects meta-analyses of woody plant biomass under elevated *vs* ambient CO_2 levels.
566 Negative estimates indicate lower CV under elevated CO_2. Models were refitted based on
567 effect sizes assuming increasing % of the effect sizes contain correlated (dependent)
568 treatment and control group data (r_{CT} = 0.8). LCI and UCI indicate the lower and upper 95%
569 confidence interval bounds. Data from Curtis, Wang [33].

% Correlated	Estimate	SE	LCI	UCI	τ^2	I^2	Q	p (Q)
0	-0.078	0.044	-0.163	0.008	0.000	0.000	85.75	0.861
20	-0.090	0.055	-0.198	0.017	0.082	35.02	141.0	0.005
40	-0.093	0.057	-0.205	0.019	0.133	55.00	191.8	<0.001
60	-0.095	0.057	-0.207	0.017	0.161	64.28	240.7	<0.001
80	-0.118	0.054	-0.225	-0.011	0.165	68.31	267.0	<0.001
100	-0.116	0.053	-0.219	-0.012	0.17	73.77	294.2	<0.001

570

571 **Table 2:** Estimates of overall effect (lnCVR) and heterogeneity from random-effects meta-
572 analyses of glycemic control in diabetics on low- *vs* high-GI diets. Negative estimates
573 indicate lower CV on a low-GI diet. Models were refitted from effect sizes assuming
574 differing strength of correlation (r_{CT}) among repeated measured from the same individuals in
575 cross-over trials. LCI and UCI indicate the lower and upper 95% confidence interval bounds.
576 Data from Brand-Miller, Petocz, Hayne, Colagiuri [34].

r_{CT}	Estimate	SE	LCI	UCI	τ^2	I^2	Q	p (Q)
0	-0.177	0.070	-0.314	-0.039	<0.001	0.006	15.88	0.321
0.3	-0.162	0.075	-0.308	-0.015	0.012	15.14	18.92	0.168
0.5	-0.151	0.080	-0.307	0.006	0.030	32.73	22.33	0.072
0.8	-0.135	0.091	-0.314	0.044	0.085	70.44	42.58	<0.001

577