Received: 17 December 2019 | Revised: 7 May 2020 | Accepted: 13 May 2020

DOI: 10.1002/jrsm.1423

RESEARCH ARTICLE

Research Synthesis Methods WILEY

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

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

[1]Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia

[2]School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia

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

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

Correspondence
Alistair M. Senior and Shinichi Nakagawa, Charles Perkins Centre, University of Sydney, Sydney, New South Wales 2006, Australia.
Email: alistair.senior@sydney.edu.au (A. M. S.) and s.nakagawa@unsw.edu.au (S. N.)

Funding information
Australian Research Council, Grant/Award Numbers: DE180101520, DP180100818; ARC Discovery Grant, Grant/Award Number: DP180100818; Australian Research Council Discovery Early Career Researcher Award, Grant/Award Number: DE180101520

Meta-analyses are often used to estimate the relative average values of a quantitative outcome in two groups (eg, control and experimental groups). However, they may also examine the relative variability (variance) of those groups. For such comparisons, two relatively new effect size statistics, the log-transformed "variability ratio" (the ratio of two standard deviations; lnVR) and the log-transformed "coefficients of variation 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 propose new estimators for lnCVR and lnVR, including for when the two groups are dependent (eg, cross-over and pre-test-post-test designs). Through simulation, we evaluated the bias of these estimators and make recommendations accordingly. We use the methods to demonstrate that: (a) lifestyle interventions have a heterogenizing effect on gestational weight gain in obese women and (b) low-glycemic index (GI) diets have a homogenizing effect on glycemic control in diabetics. We also find that the degree to which dependence among samples is accounted for can impact parameters such as τ^2 (ie, the between-study variance) and I^2 (ie, the proportion of the total variability due to between-study variance), and even the overall effect, and associated 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.

KEYWORDS
effect-size, paired design, sampling, Taylor's law, variance, variance cross-over design

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 (eg, different sexes). The standardized mean difference (SMD; also known as Cohen's *d* and its derivatives), which is the difference between group means divided by the within-study variability, is a commonly used effect size measure.[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 measure that can also be used to compare two group means is the logarithm of their ratio. This effect size measure is known as the ratio of means in medicine[3] (despite the fact it is log transformed) 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; 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 widely used in the medical and social sciences.[8] Although it should be noted that for many applications in these fields, a standardized effect size is not needed, because data tend to be reported in common units.

Two groups may not only differ in terms of their means but also their variances.[9,10] Experimental treatments may directly increase or decrease the total amount of variance in a system due to interindividual variability in response. In addition, many biological systems also appear to display a mean-variance relationship[11-13]; 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 (note Taylor's law refers to that derived by an ecologist, Taylor, as opposed to the mathematician that derived Taylor expansions or Taylor's theorem, which will be used below). This "law" has been widely observed and states that as mean population density increases, variance in population density also increases.[14,15] Where mean-variance relationships are present, a treatment may indirectly cause groups to have differing variances by altering the mean.

Highlights

- Meta-analyses typically focus on the difference between the average of two groups.
- Recently developed effect sizes now allow the user to focus on differences in the variability.
- We review existing, then suggest and test new effect sizes for meta-analysis of variability.
- We make recommendations and demonstrate the application of these new methods to worked examples from health sciences.
- Our proposed effect sizes will allow the meta-analyst to assess the difference between the variability within two groups with minimal bias.
- These new methods readily integrate with standard meta-analytic models and require no additional data than that typically required for meta-analysis of the average.

Nakagawa et al[16] proposed a number of methods that allow the user to test for differences in the variance of groups meta-analytically (for related methods, see References 10 and 17). Among the methods proposed, there are two effect sizes that readily integrate with standard contrast-based meta-analytic models.[18,19] Those two effect sizes are: (a) the logarithm of the ratio of the standard deviations (SDs), named log "variability ratio" (lnVR) and (b) the logarithm of the ratio of the coefficients of variation (CV), termed the log "CV ratio" (lnCVR). Of the two, lnCVR is perhaps the more useful measure where a mean-variance relationship (eg, Taylor's law) is likely to exist. Nakagawa et al highlight that meta-analyzing variation may be applied to completely novel datasets, but it can also be used to provide fresh insights into the topics on which a meta-analysis of means was already conducted.[16] Indeed, lnCVR has already been applied in such diverse fields as ecology,[20] evolution,[21] agriculture,[22] neuroscience,[23,24] health,[19] and the social sciences.[25] 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 article are threefold. First, we review existing and propose new estimators for lnVR and lnCVR and associated sampling error variances. These include estimators of the sampling variance when the two groups (treatment and control) are not independent (eg, in

cross-over trials or in paired, single-subject, or pre-test-post-test designs). Second, we conduct a simulation study to investigate the performance of these novel estimators. Finally, we present two case studies using these techniques and illustrate the importance of accounting for dependence between the two treatment groups in the estimation of sampling variation and other heterogeneity parameters (eg, τ^2, the between-study variance, and I^2[26]).

2 | METHODS

2.1 | Point estimators when groups are independent

Let $x_T \sim N(\mu_T, \sigma_T)$ and $x_C \sim N(\mu_C, \sigma_C)$ denote normally distributed random variables with true means (ie, expected values) given by μ_T and μ_C and true SDs σ_T and σ_C. For independent random samples based on these variables (eg, representing some outcome 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 the respective sample means and s_T and s_C the corresponding SDs for the two groups. Then comparisons among the means, variances, and CV for two groups can be made using the lnRR, lnVR, and lnCVR effect size measures, respectively. "Naïve" estimators of these effect statistics are:

$$\ln RR_{IND1} = \ln\left(\frac{\bar{x}_T}{\bar{x}_C}\right), \quad (1)$$

$$\ln VR_{IND1} = \ln\left(\frac{s_T}{s_C}\right), \quad (2)$$

$$\ln CVR_{IND1} = \ln\left(\frac{CV_T}{CV_C}\right), \quad (3)$$

where ln denotes the natural logarithm and $CV_T = s_T/\bar{x}_T$ and $CV_C = s_C/\bar{x}_C$ denote the CV in the treatment and control group, respectively.

Although 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.[16,27,28] For the lnRR, Lajeunesse[27] demonstrated such a correction is important to obtain unbiased estimation especially when sample size is small;

$$\ln RR_{IND2} = \ln\left(\frac{\bar{x}_T}{\bar{x}_C}\right) + \frac{1}{2}\left(\frac{s_T^2}{n_T\bar{x}_T^2} - \frac{s_C^2}{n_C\bar{x}_C^2}\right). \quad (4)$$

Similarly, for the lnVR, Nakagawa et al[16] proposed:

$$\ln VR_{IND2} = \ln\left(\frac{s_T}{s_C}\right) + \frac{1}{2}\left(\frac{1}{n_T-1} - \frac{1}{n_C-1}\right). \quad (5)$$

Combining $\ln RR_2$ and $\ln VR_2$, one obtains:

$$\ln CVR_{IND2} = \ln\left(\frac{CV_T}{CV_C}\right) + \frac{1}{2}\left(\frac{1}{n_T-1} - \frac{1}{n_C-1}\right) + \frac{1}{2}\left(\frac{s_C^2}{n_C\bar{x}_C^2} - \frac{s_T^2}{n_T\bar{x}_T^2}\right). \quad (6)$$

Nakagawa et al[16] originally suggested an estimator of lnCVR that missed the bias correction pertaining to lnRR (ie, $\frac{1}{2}\left(\frac{s_C^2}{n_C\bar{x}_C^2} - \frac{s_T^2}{n_T\bar{x}_T^2}\right)$). We also note here that it has been proposed that $\left(1 + \frac{1}{4n}\right)$CV acts as a "rough" bias correction for the CV (eg, [29]). From this, one could calculate the lnCVR as difference between the logarithm of the "roughly corrected" CV of each group. However, this estimator is not recommended here, and it does not perform well (see Supporting Information S1, Text S1).

2.2 | 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. However, in cross-over studies, $n_T = n_C \equiv n$. Therefore, it is useful to redefine the effect size estimators using a sample size (n) that is common to both groups. For cases of dependency, we have the naïve estimators, $\ln RR_{DEP1}$, $\ln VR_{DEP1}$, and $\ln CVR_{DEP1}$, which are identical to their independent counterparts [Equations (1)-(3)]. We can also rewrite the independent estimators based the second-order Taylor expansion for dependent cases as:

$$\ln RR_{DEP2} = \ln\left(\frac{\bar{x}_T}{\bar{x}_C}\right) + \frac{1}{2}\left(\frac{s_T^2}{n\bar{x}_T^2} - \frac{s_C^2}{n\bar{x}_C^2}\right), \quad (7)$$

$$\ln VR_{DEP2} = \ln\left(\frac{s_T}{s_C}\right) + \frac{1}{2}\left(\frac{1}{n_T-1} - \frac{1}{n_C-1}\right) = \ln\left(\frac{s_T}{s_C}\right), \quad (8)$$

$$\ln\mathrm{CVR}_{\mathrm{DEP2}} = \ln\left(\frac{\mathrm{CV_T}}{\mathrm{CV_C}}\right) + \frac{1}{2}\left(\frac{s_\mathrm{C}^2}{n\bar{x}_\mathrm{C}^2} - \frac{s_\mathrm{T}^2}{n\bar{x}_\mathrm{T}^2}\right). \quad (9)$$

It is worth highlighting that $n_\mathrm{T} = n_\mathrm{C} \equiv n$ holds unless dropouts (ie, missing response data from some individuals) are included in a pre-post design. In cases where dropouts have been included, we recommend that the sample size in both conditions (n) is assumed to be the sample size post dropouts (n_{post}). This is because the correlation between pre and post-treatment measurements can only be calculated based on the n complete samples, which assumes $n_\mathrm{T} = n_\mathrm{C}$ (see the next section).

2.3 | Dispersion estimators when the two groups are independent

The original estimators of the sampling (error) variance for lnRR[4] and lnVR[16] are based on the first-order Taylor expansion; they are, respectively,

$$s^2_{\mathrm{IND1}}(\ln\mathrm{RR}) = \frac{s_\mathrm{C}^2}{n_\mathrm{C}\bar{x}_\mathrm{C}^2} + \frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2}, \quad (10)$$

$$s^2_{\mathrm{IND1}}(\ln\mathrm{VR}) = \frac{1}{2}\left(\frac{1}{n_\mathrm{C}-1} + \frac{1}{n_\mathrm{T}-1}\right), \quad (11)$$

Based on these, for lnCVR, Nakagawa et al[16] proposed:

$$s^2_{\mathrm{IND1}}(\ln\mathrm{CVR}) = \frac{s_\mathrm{C}^2}{n_\mathrm{c}\bar{x}_\mathrm{C}^2} + \frac{1}{2(n_\mathrm{c}-1)} - 2\rho\sqrt{\frac{s_\mathrm{C}^2}{n_\mathrm{c}\bar{x}_\mathrm{C}^2}\frac{1}{2(n_\mathrm{c}-1)}} + \frac{s_T^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2} + \frac{1}{2(n_\mathrm{T}-1)} - 2\rho\sqrt{\frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2}\frac{1}{2(n_\mathrm{T}-1)}}, \quad (12)$$

where ρ is the correlation between the log mean and log SD. It was 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.[16] However, in doing so, one risks conflating within- and between-study correlation (ie, the correlation in the bivariate sampling distribution of the sample mean and sample SD could be very different to the correlation of the true means and SDs across studies). In fact, for observations that come from an underlying population distribution that is symmetric (eg, a normal distribution), the sample mean and variance are uncorrelated.[30] Thus, where $\rho = 0$, the equation above simplifies to:

$$s^2_{\mathrm{IND1}}(\ln\mathrm{CVR}) = \frac{s_\mathrm{C}^2}{n_\mathrm{C}\bar{x}_\mathrm{C}^2} + \frac{1}{2(n_\mathrm{C}-1)} + \frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2} + \frac{1}{2(n_\mathrm{T}-1)}. \quad (13)$$

As a better estimator for the sampling variance of lnRR, Lajeunesse[27] derived and tested the following sampling variance based on the second-order Taylor expansion:

$$s^2_{\mathrm{IND2}}(\ln\mathrm{RR}) = \frac{s_\mathrm{C}^2}{n_\mathrm{C}\bar{x}_\mathrm{C}^2} + \frac{s_\mathrm{C}^4}{2n_\mathrm{C}^2\bar{x}_\mathrm{C}^4} + \frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2} + \frac{s_\mathrm{T}^4}{2n_\mathrm{T}^2\bar{x}_\mathrm{T}^4}. \quad (14)$$

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

$$s^2_{\mathrm{IND2}}(\ln\mathrm{VR}) = \frac{1}{2}\left(\frac{1}{n_\mathrm{C}-1} + \frac{1}{(n_\mathrm{C}-1)^2} + \frac{1}{n_\mathrm{T}-1} + \frac{1}{(n_\mathrm{T}-1)^2}\right) = \frac{1}{2}\left(\frac{n_\mathrm{C}}{(n_\mathrm{C}-1)^2} + \frac{n_{TT}}{(n_\mathrm{T}-1)^2}\right). \quad (15)$$

Accordingly, the complete estimator of the sampling variance for lnCVR, based on $s^2(\mathrm{lnRR}_{\mathrm{IND2}})$ and $s^2(\mathrm{lnVR}_{\mathrm{IND2}})$ is:

$$s^2_{\mathrm{IND2}}(\ln\mathrm{CVR}) = \frac{s_\mathrm{C}^2}{n_\mathrm{C}\bar{x}_\mathrm{C}^2} + \frac{s_\mathrm{C}^4}{2n_\mathrm{C}^2\bar{x}_\mathrm{C}^4} + \frac{n_\mathrm{C}}{(n_\mathrm{C}-1)^2} + \frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2} + \frac{s_T^4}{2n_\mathrm{T}^2\bar{x}_\mathrm{T}^4} + \frac{n_\mathrm{T}}{(n_\mathrm{T}-1)^2}. \quad (16)$$

In the Supporting Information, 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[31] (see Supporting Information S1, Text S2).

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 (ie, cross-correlation[32]). Based on the first-order Taylor expansion, the sampling variance for dependent lnRR is:

$$s^2_{\mathrm{DEP1}}(\ln\mathrm{RR}) = \frac{s_\mathrm{C}^2}{n_\mathrm{C}\bar{x}_\mathrm{C}^2} + \frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2} - 2r_{\mathrm{CT}}\sqrt{\frac{s_\mathrm{C}^2}{n_\mathrm{C}\bar{x}_\mathrm{C}^2}}\sqrt{\frac{s_\mathrm{T}^2}{n_\mathrm{T}\bar{x}_\mathrm{T}^2}}, \quad (17)$$

where r_{CT} is a cross-condition correlation value estimated from the two sets of measurements on the same replicate when they are under the control and treatment conditions.[33] As discussed above for dependent studies $n_T = n_C \equiv n$, meaning $s^2_{DEP1}(\text{lnRR})$ simplifies to:

$$s^2_{DEP1}(\text{lnRR}) = \frac{s^2_C}{n\bar{x}^2_C} + \frac{s^2_T}{n\bar{x}^2_T} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T}. \quad (18)$$

If based on the second-order Taylor expansion,[27] the estimator of the sampling variance for lnRR is:

$$s^2_{DEP2}(\text{lnRR}) = \frac{s^2_C}{n\bar{x}^2_C} + \frac{s^2_T}{n\bar{x}^2_T} + \frac{s^4_C}{2n^2\bar{x}^4_C} + \frac{s^4_T}{2n^2\bar{x}^4_T} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T} + r^2_{CT}\frac{s^2_C s^2_T\left(\bar{x}^4_C + \bar{x}^4_T\right)}{2n^2\bar{x}^4_C\bar{x}^4_T}. \quad (19)$$

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

$$s^2_{DEP1}(\text{lnVR}) = \frac{1}{2}\left(\frac{1}{(n_C-1)} + \frac{1}{(n_T-1)}\right) - r^2_{CT}\sqrt{\frac{1}{(n_C-1)}}\sqrt{\frac{1}{(n_T-1)}}, \quad (20)$$

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

$$s^2_{DEP1}(\text{lnVR}) = \frac{1-r^2_{CT}}{n-1}. \quad (21)$$

Based on the second-order Taylor expansion, we have the sampling variance for dependent cases of lnVR as:

$$s^2_{DEP2}(\text{lnVR}) = \frac{n}{(n-1)^2} - r^2_{CT}\frac{1}{n-1} + r^4_{CT}\frac{s^8_C + s^8_T}{2(n-1)^2 s^4_C s^4_T}. \quad (22)$$

From the sampling variances for lnRR and lnVR, we have the sampling variance for lnCVR with the first- and second-order Taylor expansions as:

$$s^2_{DEP1}(\text{lnCVR}) = \frac{s^2_C}{n\bar{x}^2_C} + \frac{s^2_T}{n\bar{x}^2_T} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T} + \frac{1}{n-1} - r^2_{CT}\frac{1}{n-1}, \quad (23)$$

$$s^2_{DEP2}(\text{lnCVR}) = \frac{s^2_C}{n\bar{x}^2_C} + \frac{s^2_T}{n\bar{x}^2_T} + \frac{s^4_C}{2n^2\bar{x}^4_C} + \frac{s^4_T}{2n^2\bar{x}^4_T} - r_{CT}\frac{2s_C s_T}{n\bar{x}_C\bar{x}_T} + r^2_{CT}\frac{s^2_C s^2_T\left(\bar{x}^4_C + \bar{x}^4_T\right)}{2n^2\bar{x}^4_C\bar{x}^4_T} + \frac{n}{(n-1)^2} - r^2_{CT}\frac{1}{n-1} + r^4_{CT}\frac{s^8_C + s^8_T}{2(n-1)^2 s^4_C s^4_T}. \quad (24)$$

Note that, where r is positive the estimated sample variance for a dependent estimator will be smaller than its independent equivalent, but that as r shrinks to 0, the dependent case converges on the independent; e.g. assuming $n_C = n_T$, where $r > 0$, $s^2_{DEP1}(\text{lnCVR}) < s^2_{IND1}(\text{lnCVR})$, but where $r = 0$, $s^2_{DEP1}(\text{lnCVR}) = s^2_{IND1}(\text{lnCVR})$.

3 | SIMULATION

3.1 | Simulation study design

We simulated a two-group experiment/trial, where each group is based on n_T and n_C random samples drawn from populations under experimental treatment and control conditions. The treatment and control populations have means μ_T and μ_C and SDs σ_T and σ_C, respectively. The ith sample in the treatment and control groups, y_{Ti} and y_{Cj} ($i = 1 \ldots \max[n_C, n_T]$) was drawn from a bivariate normal distribution as follows:

$$\begin{pmatrix} y_{Ti} \\ y_{Ci} \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_T \\ \mu_C \end{bmatrix}, \begin{bmatrix} \sigma^2_T & \rho_{CT}\sigma_T\sigma_C \\ \rho_{CT}\sigma_C\sigma_T & \sigma^2_C \end{bmatrix}\right), \quad (25)$$

where $\begin{bmatrix} \mu_T \\ \mu_C \end{bmatrix}$ are the population means of the two groups, $\begin{bmatrix} \sigma^2_T & \rho_{CT}\sigma_T\sigma_C \\ \rho_{CT}\sigma_C\sigma_T & \sigma^2_C \end{bmatrix}$ is a variance-covariance matrix specifying the variances of the two groups with ρ_{CT} giving the degree of correlation among the ith samples in the two groups and all other parameters are as above. Where $\rho_{CT} \neq 0$, the ith data in the two groups are correlated (ie, dependent or paired samples as in a cross-over design). We explored n_C = 8, 16, and 42, with $n_C = n_T$. For the independent case (ie, $\rho_{CT} = 0$), we also explored $n_C < n_T$, and here we simulated $\max(n_C, n_T)$ values in each group before randomly deleting data to achieve the desired sample sizes.

In all simulations, μ_C = 100 and σ_C = 20, which across the parameters tested ensures positive sample means (required for log transformation). We explored values of μ_T ranging between $\mu_C \times e^{-0.5}$ 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 $\ln(\mu_T/\mu_C)$ and $\ln(\sigma_T/\sigma_C)$ is between −0.5 and 0.5. All combinations were explored and where $\ln(\mu_T/\mu_C) = \ln(\sigma_T/\sigma_C)$, the coefficient of variation (CV) of the two groups will be identical. We also explored $\rho_{CT} = 0$

and $\rho_{CT} = 0.8$. For each set of parameters, we simulated 100 000 experiments, which we found gave smooth visualization of any trends over the parameter space.

Here, we focus on the performance of estimators of lnCVR, because: (a) the bias adjustments to lnCVR are a composite of those described for lnRR and lnVR, meaning it simultaneously tests both and (b) the described adjustments to lnRR have already been tested extensively and found to perform well.[27] Based on the sample means and SDs of each simulated experiment, we calculated $\text{lnCVR}_{\text{IND1}}$ and $\text{lnCVR}_{\text{IND2}}$ for independent cases ($\rho_{CT} = 0$) and $\text{lnCVR}_{\text{DEP1}}$ and $\text{lnCVR}_{\text{DEP2}}$ for dependent cases ($\rho_{CT} \neq 0$). We also calculated the sampling variance estimators $s^2_{\text{IND1}}(\text{lnCVR})$ and $s^2_{\text{IND2}}(\text{lnCVR})$, where $\rho_{CT} = 0$, and $s^2_{\text{DEP1}}(\text{lnCVR})$ and $s^2_{\text{DEP2}}(\text{lnCVR})$ where $\rho_{CT} \neq 0$. We calculated bias in the lth estimator (l = IND1, IND2, DEP1, and DEP2) as:

$$\text{bias}[\text{lnCVR}_l] = \frac{1}{K}\sum_{k=1}^{K}\text{lnCVR}_{lk} - \ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right), \quad (26)$$

where k is the kth simulated value (k = 1... K; K = 100 000) of lnCVR_l. This bias can be interpreted as the mean deviation of the lth estimator of lnCVR from the true population value. We calculated relative bias in sampling variance estimator l as:

$$\text{bias}\left[s^2_l(\text{lnCVR})\right] = \frac{s^2_l(\text{lnCVR}) - \theta^2}{\theta^2} \times 100, \quad (27)$$

where $s^2_l(\text{lnCVR})$ is the value of the lth sampling variance based on the simulated population statistics and sample sizes, and θ is the SD among K simulated effect sizes quantified using the estimator determined to have minimal bias [as determined by Equation (26)]. This bias can be interpreted as the percentage by which the sampling variance estimator deviates from the true value (ie, 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 the lth sampling variance $s^2_l(\text{lnCVR})$, 95% CIs were constructed as:

$$95\%\text{CI} = \text{lnCVR} \pm z_{0.975}s_l(\text{lnCVR}) \quad (28)$$

where lnCVR is the estimated effect size for the simulated sample quantified using the estimator with minimal bias, $s_l(\text{lnCVR})$ an estimate of the SE (the square root of $s^2_l(\text{lnCVR})$), and $z_{0.975}$ is the function of the 0.975th quantile of a z distribution (~1.96). Simulations and analyses were performed in R v3.5.1[34] and used the "mvrnorm" function in the *MASS* package.[35]

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 effects 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 ($\ln\left(\frac{\sigma_T/\mu_T}{\sigma_C/\mu_C}\right) < 0$). $\text{lnCVR}_{\text{IND1}}$ overestimates positive effects and slightly underestimates negative effects, with bias being most profound where the sample size is small. $\text{lnCVR}_{\text{IND2}}$, conversely, 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. $\text{lnCVR}_{\text{IND1}}$ showed the greatest upward bias, especially where the sample size was small. On the other hand, $\text{lnCVR}_{\text{IND2}}$ performed with only very minor upward bias, which all but disappeared for larger sample sizes. Given that $\text{lnCVR}_{\text{IND2}}$ was determined to be the most accurate estimator of the effect, we proceeded to explore how $\text{lnCVR}_{\text{IND2}}$ performed in conjunction with different estimators of sampling variance.

The first sampling variance estimator $s^2_{\text{IND1}}(\text{lnCVR})$ underestimated the variance among simulated values of $\text{lnCVR}_{\text{IND2}}$, and where the sample size was small, this underestimate was by around 10% (Figure 3). Biases for $s^2_{\text{IND2}}(\text{lnCVR})$ were minimal, although there was some very slight upward bias for small sample sizes and large positive effects (Figure 3). The coverage of 95% CIs for $s^2_{\text{IND1}}(\text{lnCVR})$ and $s^2_{\text{IND2}}(\text{lnCVR})$ (paired with $\text{lnCVR}_{\text{IND2}}$) are shown in Figure 4. $s^2_{\text{IND1}}(\text{lnCVR})$ generated CIs that were too narrow at smaller sample sizes, whereas again $s^2_{\text{IND2}}(\text{lnCVR})$ performed with little bias. At larger sample sizes, coverage was much closer to the nominal level (Figure 4), although $s^2_{\text{IND2}}(\text{lnCVR})$ still performed more accurately. The same patterns of performance were observed for the case where $n_C < n_T$ (Supporting Information Figures S1 and S2).

For the case where treatment and control samples were dependent on one another ($\rho_{CT} = 0.8$), $\text{lnCVR}_{\text{DEP2}}$ out-performed $\text{lnCVR}_{\text{DEP1}}$, with a pattern identical to that in Figure 1 (Supporting Information Figure S3). With regards estimators for dependent sampling variances, $s^2_{\text{DEP1}}(\text{lnCVR})$ underestimated the variance where as $s^2_{\text{DEP2}}(\text{lnCVR})$ overestimated the variance (Figure 5). These biases were within a reasonable range for larger samples, but were severe for small samples.

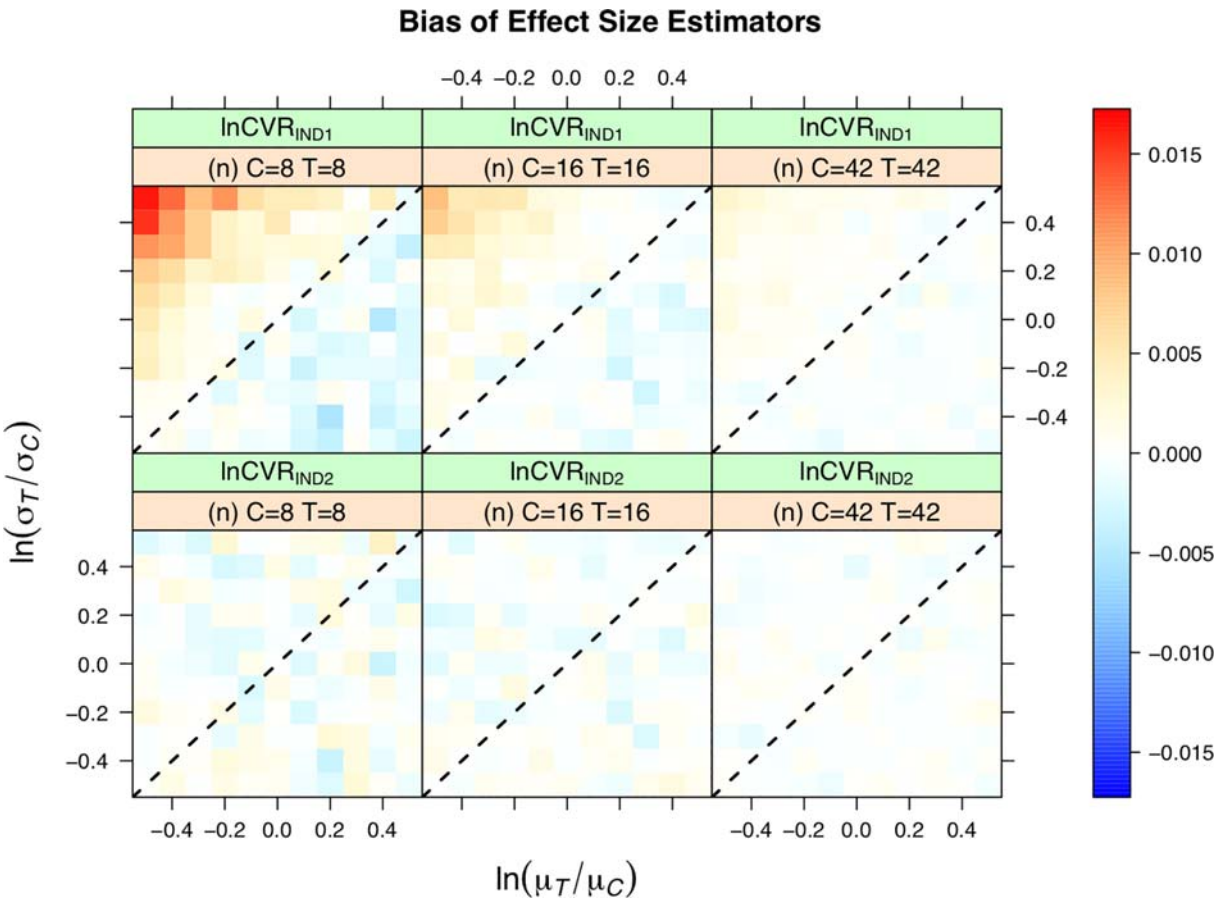

FIGURE 1 Bias in effect size estimators of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis), and sample size (balanced) for the case of independent treatment and control group data ($\rho_{CT} = 0$). Black dashed line indicates no effect (ie, lnCVR = 0). lnCVR, log-transformed coefficients of variation ratio [Colour figure can be viewed at wileyonlinelibrary.com]

$s^2_{DEP2}(\text{lnCVR})$ 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 $s^2_{DEP1}(\text{lnCVR})$ had a tendency to be too narrow, whereas those generated by $s^2_{DEP2}(\text{lnCVR})$ were too wide (Figure 6).

4 | WORKED EXAMPLES

We now provide two examples from the health sciences. All meta-analytic models (random-effects meta-analysis) were fitted using the "rma" function (with default settings) in *metafor*.[36]

4.1 | Example 1. *Interventions to control gestational weight gain*

Yeo et al[37] assessed the effects of lifestyle interventions (diet and/or physical exercise) on gestational weight gain in obese women. The studies were varied somewhat in the mode of intervention, and the period of pregnancy covered. Individual participants in control and intervention groups in all study designs were considered independent. Using a random-effects model they found that, based on 20 effect sizes, interventions led to a statistically significant reduction in mean gestational weight gain by on average 2.07 kg relative to a concurrent control group.[37] For both control and intervention groups, there was a positive correlation between the log sample mean and SD (Figure 7A).

We first reanalyzed the effects of dietary interventions on mean weight gain on the ratio scale via lnRR_{IND2} and sampling variance $s^2_{IND2}(\text{lnRR})$ using a random-effects model. Like Yeo et al,[37] we found that intervention groups had a significantly lower weight gain than control groups (lnRR = −0.225, LCI = −0.363 to UCI = −0.086); the median weight gain of intervention groups was 19.25% (1 − exp[−0.225]) lower than controls. We next aimed to assess how dietary interventions affect among-participant variability in gestational weight gain.

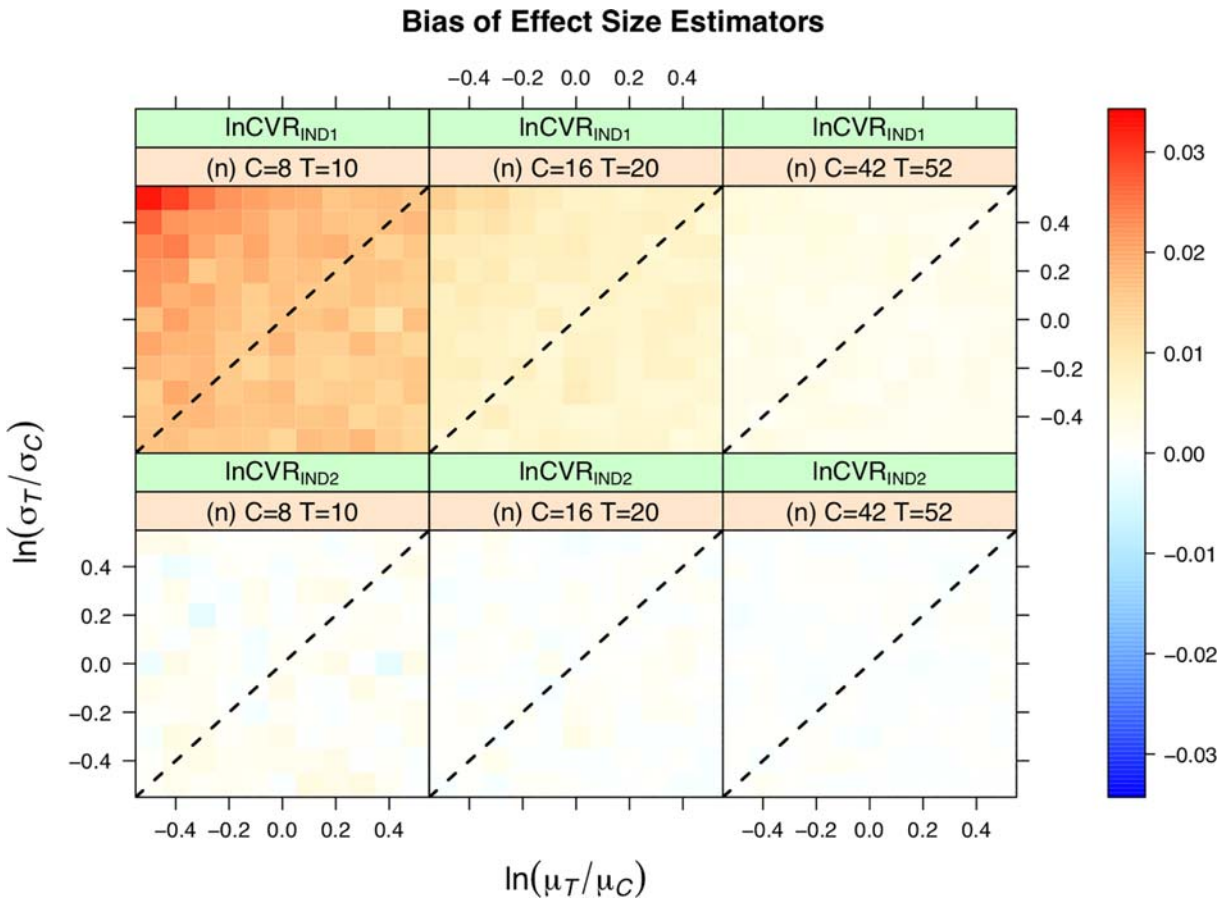

FIGURE 2 Bias in effect size estimators of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis), and sample size (unbalanced) for the case of independent treatment and control group data ($\rho_{CT} = 0$). Black dashed line indicates no effect (ie, lnCVR = 0). lnCVR, log-transformed coefficients of variation ratio [Colour figure can be viewed at wileyonlinelibrary.com]

Given that there is an apparent association between the population mean and variance (Figure 7A), it arguably makes most sense to test whether interventions generate more variation than controls after correcting for differences in the population means. Put another way, is there lower variation in the treatment group relative to control than we would expect given that they have different means. Accordingly, we assessed whether diet affects variation using $\text{lnCVR}_{\text{IND2}}$ and sampling variance $s^2_{\text{IND2}}(\text{lnCVR})$. We do note that if we were explicitly interested in the effects of interventions on variance as a statistical quantity (perhaps for questions related to statistical power), $\text{lnVR}_{\text{IND2}}$ may be more useful. We found that overall there was a statistically significant positive estimate (lnCVR = 0.245, LCI = 0.108 to UCI = 0.381), whereby interventions increased the CV in weight gain among participants by 27.76% relative to controls. There was some evidence for moderate heterogeneity in the effect of interventions on among-participant variation ($\tau^2 = 0.04$, $I^2 = 52.63\%$).

4.2 | Example 2. *Low glycemic index diets and glycemic control in diabetic subjects*

Brand-Miller et al[38] performed a meta-analysis of studies designed to test the effects of low-glycemic index (GI) diets on biomarkers of glycemic control in diabetic (types 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 et al[38] acknowledged that for those studies with a cross-over design, there will be a degree of correlation among the treatment and control condition data.

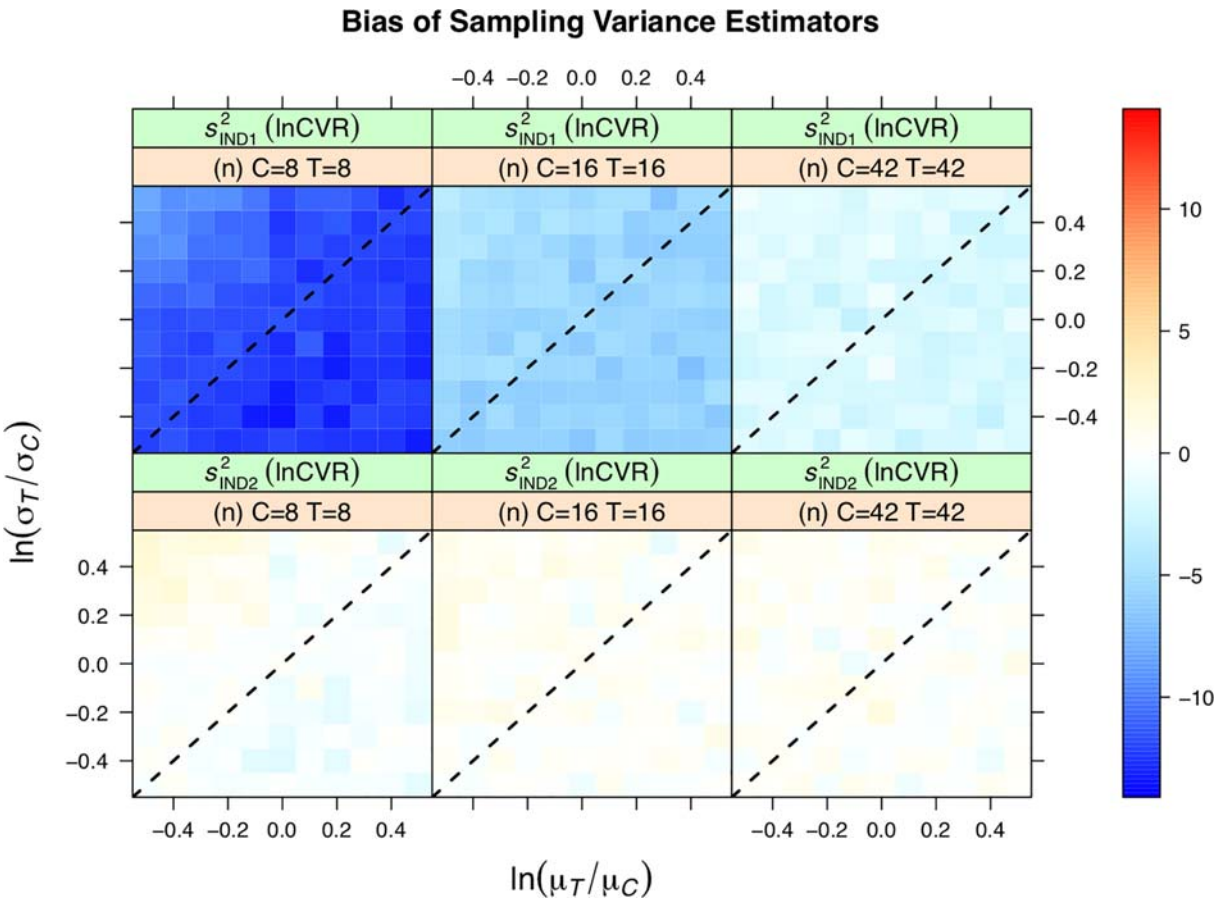

FIGURE 3 Relative bias in sampling variance estimators of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis), and sample size (balanced) for the case of independent treatment and control group data ($\rho_{CT} = 0$). Black dashed line indicates no effect (ie, lnCVR = 0). lnCVR, log-transformed coefficients of variation ratio [Colour figure can be viewed at wileyonlinelibrary.com]

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 interindividual variability in glycemic control using lnCVR. Unlike Example 1 here, there are studies that 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 $\text{lnCVR}_{\text{IND2}}$, and $s^2_{\text{IND2}}(\text{lnCVR})$. For those studies using a cross-over design, we calculated $\text{lnCVR}_{\text{DEP2}}$ and $s^2_{\text{DEP1}}(\text{lnCVR})$ 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 among 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 1. The analyses consistently estimated a negative overall effect size, suggesting that on low-GI diets the CV in biomarkers of glycemic control is on average reduced 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 1); for $r_{CT} = 0.5$, the overall effect was not statistically significant. With increasing correlation, heterogeneity also increased (Table 1). 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 1).

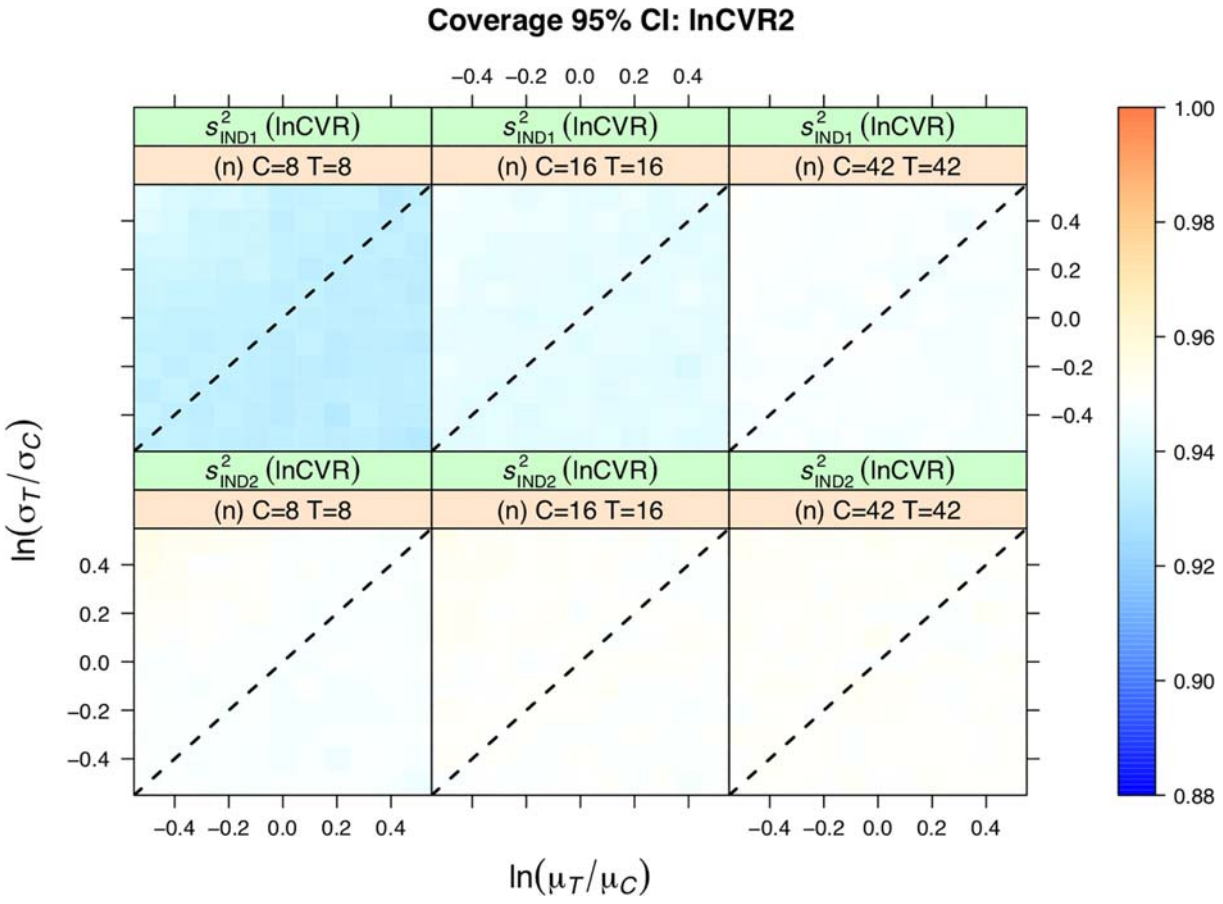

FIGURE 4 Coverage of 95% CIs based on estimators of the sampling variance of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis), and sample size (balanced) for the case of independent treatment and control group data ($\rho_{CT} = 0$). Black dashed line indicates no effect (ie, lnCVR = 0). CI, confidence interval; lnCVR, log-transformed coefficients of variation ratio [Colour figure can be viewed at wileyonlinelibrary.com]

5 | DISCUSSION AND CONCLUSIONS

We recommend that meta-analysts estimate lnCVR for independent study designs using $\text{lnCVR}_{\text{IND2}}$ [Equation (6)], and for dependent study designs, we recommend the use of the $\text{lnCVR}_{\text{DEP2}}$ [Equation (9)]. Under the simulated conditions explored, these estimators exhibited minimal bias. In contrast, "naïve" estimators displayed systematic biases, substantially overestimating large positive effects, especially when sample sizes were small. Compared with previous estimators,[16] 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.[27]

We also recommend that the sampling variance of lnCVR be estimated for independent and dependent study designs using $s^2_{\text{IND2}}(\text{lnCVR})$ [Equation (16)] and $s^2_{\text{DEP1}}(\text{lnCVR})$ [Equation (23)], respectively. 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* distribution) that are slightly too narrow. This may well 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, $s^2_{\text{DEP2}}(\text{lnCVR})$ is given in Equation (24) 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: (a) lifestyle interventions affect interindividual gestational weight gain in obese participants, and

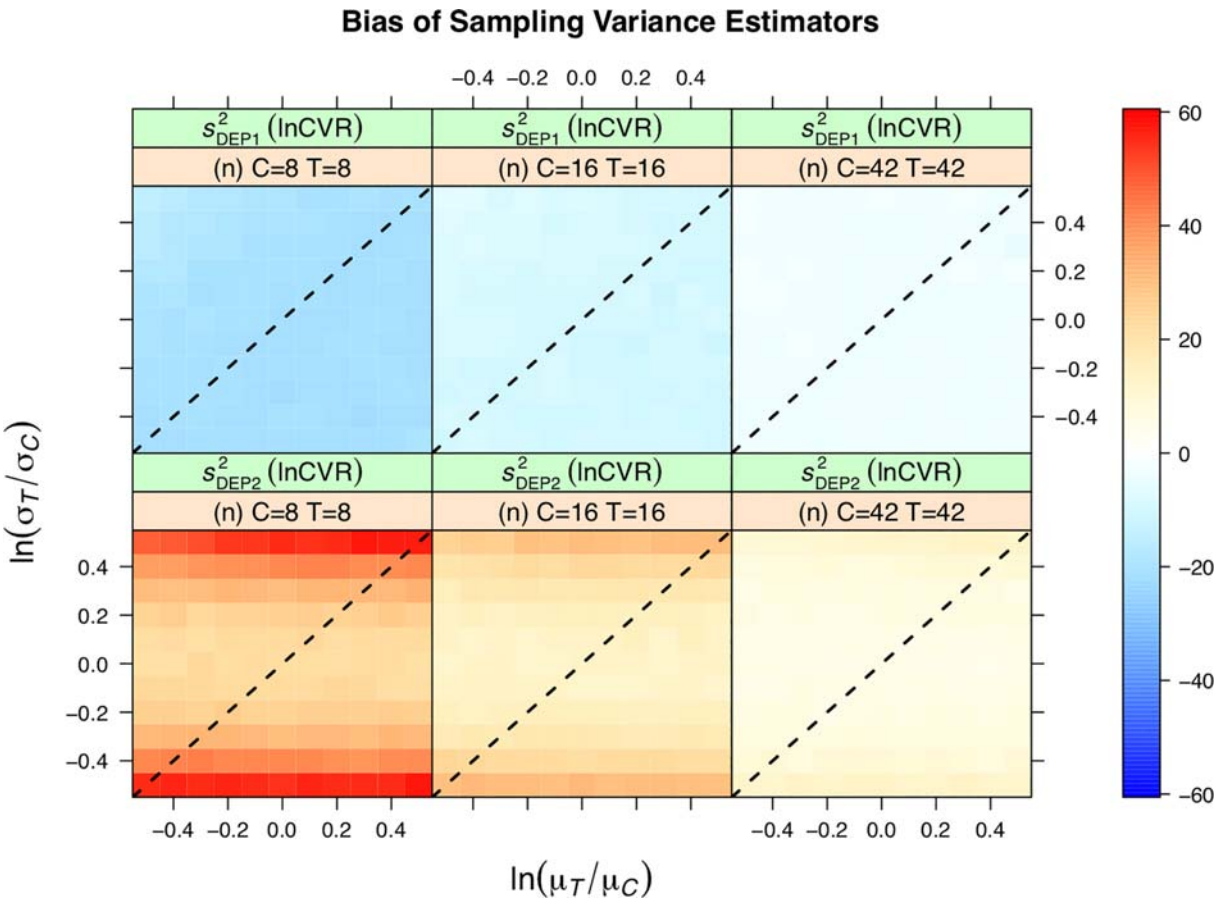

FIGURE 5 Relative bias in sampling variance estimators of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis), and sample size (balanced) for the case of dependent treatment and control group data ($\rho_{CT} = 0.8$). Black dashed line indicates no effect (ie, lnCVR = 0). lnCVR, log-transformed coefficients of variation ratio [Colour figure can be viewed at wileyonlinelibrary.com]

(b) low-GI diets alter between-individual variation in glycemic control in diabetics. In Example 1, we found that interventions reduce weight gain during gestation on average by around 19% but increase the CV in weight gain by over 27%. This is indicative that most lifestyle interventions (there was moderate heterogeneity among studies in the effect) cause heterogenization in gestational weight gain, whereby there are responders and non-responders. Clear questions for researchers working in the field are: (a) are nonresponders identifiable and (b) do these treatments actually cause excessive weight gain in a subpopulation (see also Reference 17).

In the second example of low-GI diets and glycemic control, we found that these diets have a homogenizing effect (ie, decrease the CV). This result coupled with a beneficial effect on the mean indicates that low-GI diets are likely to be beneficial for most individuals in the populations sampled. However, 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 SE). 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 affects conclusions about the consistency of the reported effects. This trend can be explained by the fact that as stronger correlations are assumed the sampling variances associated with the individual effect sizes shrink, effects are assumed to be more precise, and sampling variability becomes less able to explain the variation among the effects. Our results corroborate the points made by Becker,[32] who introduced an estimator for the sampling variance of SMD for dependent groups.

To inspire and facilitate broader adoption of the meta-analysis of variation, we briefly summarize three more examples of lnVR and/or lnCVR applied to different fields. First, applying lnCVR to nearly 200 studies, Knapp and van der Heijden[22] have shown that organic agriculture has 15% less stability in yield than conventional agriculture. Such instability is certainly a disadvantage of organic agriculture, although these drawbacks

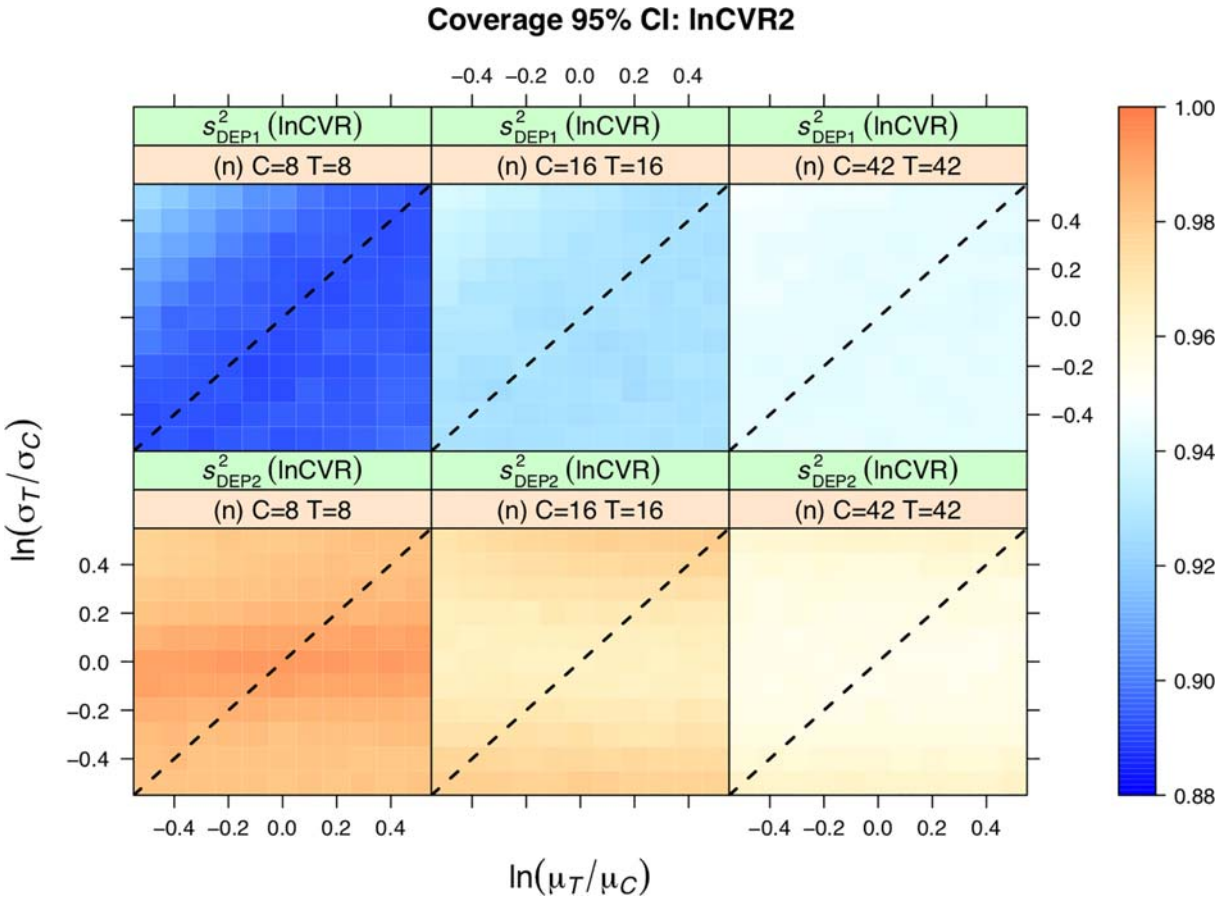

FIGURE 6 Coverage of 95% CIs based on estimators of the sampling variance of lnCVR as a function of the log ratio of population means (*x*-axis), SDs (*y*-axis), and sample size (balanced) for the case of dependent treatment and control group data ($\rho_{CT} = 0.8$). Black dashed line indicates no effect (ie, lnCVR = 0). CI, confidence interval; lnCVR, log-transformed coefficients of variation ratio [Colour figure can be viewed at wileyonlinelibrary.com]

must be balanced against the benefits of organic farming for biodiversity. Second, Brugger et al[23] tested, for the first time, the hypothesis that patients with schizophrenia exhibit heterogeneity in dopamine function by combining 65 studies with both lnVR and lnCVR. They found that schizophrenic patients have more variability in the availability of dopamine receptors and transporters than control subjects. Finally, O'Dea et al[25] have confirmed that boys' school grades are more variable than those of girls, by applying lnCVR to data from over 1.6 million students. However, contrary to conventional expectations, the variability between girls and boys was similar in the STEM subjects. Rather, the largest difference in variability was observed in languages.

Despite the relative simplicity of lnCVR, 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. 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 (eg, in questions related to power and study design) and again lnVR would be an appropriate choice. However, where mean-variance relationships exist, 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 among species/traits that exist on different scales because CV is a relative measure.[39] We note that such a practice is not necessarily required for meta-analysis, because lnVR is also a relative measure of variation. 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.

Like most pairwise effect sizes, those presented here assume the underlying data follow a normal distribution,

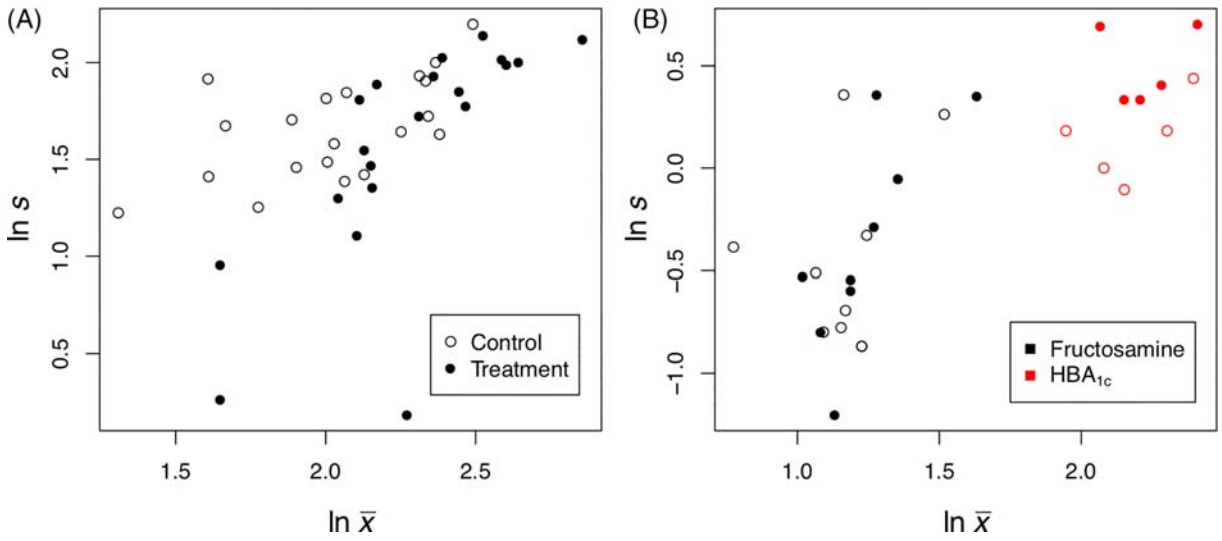

FIGURE 7 Association between log sample mean (ln $\bar{x}$) and log sample SD (ln s) for treatment (hollow points) and control (solid points) groups in different studies. Data are from: A, Yeo et al,[37] where the outcome is gestational weight gain in obese women subjected to lifestyle interventions (treatment) and control conditions (control); and B, Brand-Miller et al,[38] where the outcome is a measure of glycemia in diabetic individuals on low (treatment) vs high (control) glycemic index diets. Note in (B) measures of glycemia are either fructosamine (black points) or HbA_{1c} (red points) levels, where lower levels indicate better gylcemic control [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Estimates of overall effect (lnCVR) and heterogeneity from random-effects meta-analyses of GI control in diabetics on low- vs high-GI diets

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	.321
0.3	−0.162	0.075	−0.308	−0.015	0.012	15.14	18.92	.168
0.5	−0.151	0.080	−0.307	0.006	0.030	32.73	22.33	.072
0.8	−0.135	0.091	−0.314	0.044	0.085	70.44	42.58	<.001

Note: Negative estimates indicate lower CV on a low-GI diet. Models were refitted from effect sizes assuming differing strength of correlation (r_{CT}) among repeated measures from the same individuals in cross-over trials.
Abbreviations: CV, coefficients of variation; GI, glycemic index; LCI, lower 95% confidence interval; lnCVR, log-transformed coefficients of variation ratio; SE, standard error; UCI, upper 95% confidence interval.
Source: Data from Brand-Miller et al.[38]

which may not always be the case. For example, often biological data are log normally distributed.[40] If the user is concerned that the data underlying their effect sizes are log normally distributed, there are remedial measures that they can take (see Supporting Information S1, Text S3). Nakagawa et al[16] 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. These models also allow the user to avoid the constraint of positive-only sample means, which is a requirement for ratio-based effect sizes such as lnCVR and lnRR. Arm-based models, however, are not without their critics who argue that these methods are radical departure from established meta-analytic thinking (see Reference 18). Like other (contrast-based) effect size measures that reflect the difference between two groups (eg, SMD, lnRR, log risk/odds ratio, or the risk difference), lnCVR readily integrates with our most widespread analytical paradigms.

Finally, we finish by reiterating the point made by Nakagawa et al[16] and echoed by subsequent papers using lnCVR in different fields of study.[19-25] As we have discussed and shown, meta-analysis of variation can tackle entirely new questions and open our eyes to insights that are hidden within datasets. The data 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 (cf Reference 41). In

this regard, it is vital that meta-analytic "raw" data are made available and reusable in the spirit of open and transparent science.[42,43]

ACKNOWLEDGEMENTS

A.M.S. was supported by an Australian Research Council Discovery Early Career Researcher Award (ARC DECRA: DE180101520). S.N. 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.

CONFLICT OF INTEREST

The author reported no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in https://github.com/AlistairMcNairSenior/lnCVR_Estimators_Sim.

ORCID

Alistair M. Senior https://orcid.org/0000-0001-9805-7280

Wolfgang Viechtbauer https://orcid.org/0000-0003-3463-4063

Shinichi Nakagawa https://orcid.org/0000-0002-7765-5182

REFERENCES

1. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*. San Diego, CA: 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*. New York: 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. Prendergast LA, Staudte RG. Meta-analysis of ratios of sample variances. *Stat Med*. 2016;35(11):1780-1799.
11. Cohen JE, Xu M. Random sampling of skewed distributions implies Taylor's power law of fluctuation scaling. *Proc Natl Acad Sci U S A*. 2015;112(25):7749-7754.
12. 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 U S A*. 2012;109(39):15829-15834.
13. Nakagawa S, Schielzeth H. The mean strikes back: mean-variance relationships and heteroscedasticity. *Trends Ecol Evol*. 2012;27(9):474-475.
14. Taylor LR. Aggregation, variance and the mean. *Nature*. 1961;189(4766):732-735.
15. Bartlett MS. Some notes on insecticide tests in the laboratory and in the field. *J R Stat Soc*. 1936;3(2):185-194.
16. 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.
17. Cortés J, González JA, Medina MN, et al. Does evidence support the high expectations placed in precision medicine? A bibliographic review. *F1000Res*. 2018;7:30.
18. Dias S, Ades AE. Absolute or relative effects? Arm-based synthesis of trial data. *Res Syn Methods*. 2016;7:23-28.
19. 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.
20. 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.
21. Janicke T, Häderer IK, Lajeunesse MJ, Anthes N. Darwinian sex roles confirmed across the animal kingdom. *Sci Adv*. 2016;2(2):e1500983.
22. Knapp S, van der Heijden MGA. A global meta-analysis of yield stability in organic and conservation agriculture. *Nat Commun*. 2018;9(1):3632.
23. Brugger SP, Angelescu I, Abi-Dargham A, Mizrahi R, Shahrezaei V, Howes OD. Heterogeneity of striatal dopamine function in schizophrenia: meta-analysis of variance. *Biol Psychiatry*. 2020;87(3):215-224.
24. Mizuno Y, McCutcheon RA, Brugger SP, Howes OD. Heterogeneity and efficacy of antipsychotic treatment for schizophrenia with or without treatment resistance: a meta-analysis. *Neuropsychopharmacology*. 2020;45(4):622-631.
25. 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.
26. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558.
27. Lajeunesse MJ. Bias and correction for the log response ratio in ecological meta-analysis. *Ecology*. 2015;96(8):2056-2063.
28. Raudenbush SW, Bryk AS. Examining correlates of diversity. *J Edu Stat*. 1987;12:241-269.
29. Sokal R, Rohlf FJ. *Biometry*. 3rd ed. New York: Freeman; 1995.
30. Zhang L. Sample mean and sample variance. *Am Stat*. 2007;6(2):159-160.
31. Noble DWA, Lagisz M, O'Dea RE, Nakagawa S. Non-independence and sensitivity analyses in ecological and evolutionary meta-analyses. *Mol Ecol*. 2017;26(9):2410-2425.

32. Becker BJ. Synthesizing standardized mean-change measures. *Br J Math Stat Psychol*. 1988;41(2):257-278.
33. Lajeunesse MJ. On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology*. 2011; 92(11):2049-2055.
34. R: A language and environemnt for statistical computing [computer program]. Version 3.5.1. http://www.r-project.org.2018. Accessed March 25, 2020.
35. Venables WN, Ripley BD. *Modern Applied Statistics with S*. 4th ed. New York: Springer; 2002.
36. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36:1-48.
37. Yeo S, Walker JS, Caughey MC, Ferraro AM, Asafu-Adjei JK. What characteristics of nutrition and physical activity interventions are key to effectively reducing weight gain in obese or overweight pregnant women? A systematic review and meta-analysis. *Obes Rev*. 2017;18(4):385-399.
38. 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.
39. Van Valen L. The statistics of variation. *Evol Theory*. 1978;4:433-443.
40. Limpert E, Stahel WA, Abbt M. Log-normal distributions across the sciences: keys and clues. *Bioscience*. 2001;51(5):341-352.
41. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature*. 2018;8 (555):175-182.
42. Nosek BA, Alter G, Banks GC, et al. Promoting an open research culture. *Science*. 2015;348(6242):1422-1425.
43. 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.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Senior AM, Viechtbauer W, Nakagawa S. Revisiting and expanding the meta-analysis of variation: The log coefficient of variation ratio. *Res Syn Meth*. 2020; 11:553–567. https://doi.org/10.1002/jrsm.1423