Consulting Forum Member Request for Peer Review of the following Claim:

In Random-effects meta-analysis, the mainstream methods estimate the overall effect as

 $\hat{θ}$ = ΣWj $\hat{θ}$j (1)

where the Wj are the study weights (summing to one), and the $\hat{θ}$j are the study-specific estimates of effect for study j, j=1,2,…,M. The studies are presumed to be independent. The key question that will underpin the validity of this mainstream approach is whether the weights Wj are (A) constants (or even nearly constants) or (B) seriously random variables.

The seriousness of this question may be reflected that over 20,000 PUBMED titles, in 2019 alone, used the mainstream methods (1). If the answer is (B) then it will follow below that the overall estimate may be seriously biased and that the universally used asymptotic variance formula for $\hat{θ}$ per (4) below, is incorrect. Meta-analysis stands at the APEX of most evidence pyramids, as you can see by a Google Search. If (B) is correct, this totally undermines a 44-year history, and may call into doubt major health policy decisions that relied in a major way upon meta-analysis.

I am asking you to review the content below, whose mathematical level can be understood by graduate students in statistics or biostatistics, and vote on whether whether (A) is correct or (B) is correct. **No prior knowledge of Meta-analysis is needed to review this. It is self-contained.**

Thanks,

Jon Shuster

1. **Demonstration that the Weights are seriously random variables**

First, we shall illustrate the compelling validity problems with a simple example to demonstrate that weights must be considered as random variables, not constants. This is a fatal flaw to the mainstream weighted methods, ultimately rendering them as being potentially biased and having incorrect variance estimates of the main outcome parameter.

Suppose we have three studies in our meta-analysis (Abbott, Barnes, and Cole) with respective weights (0.50, 0.30, 0.20). They are presently unindexed. Before we give the data to the statistician, we will order the studies randomly where each of the six orderings is equally likely. The statistician has no basis for any complaint, since her/his ultimate weighted analysis is not affected by the random ordering. If Wj is the weight assigned to Study j, it is conditionally uniform over the weights (0.50, 0.30, 0.20) given the three unassigned weights. Hence the conditional mean and variance of the weights are 0.3333 (1/3) and 0.0156 respectively. The conditional coefficient of variation 100(SD/Mean)% is a substantial 37%. If the weights are supposed to be constants, this conditional coefficient of variation should be zero. If you grasp this, it will follow that you reject weighted methods.

1. ***Demonstration that Mainstream Weighted Methods (Effects-at-Random) are Invalid.***

In this subsection, we prove that the mainstream weighted methods are invalid. Optimization of weights inversely proportional to the variance of the individual study estimates is not only invalid but is optimizing a metric that is not the true variance for the overall weighted estimator! Note that we are forgiving any sampling error made in creating weights, and every method of weighting (other than uniform) will produce the same deficiencies. (Note that we do not advocate any weighted method, including uniform.)

 *Framework*: Suppose we have M independent studies in our meta-analysis, and these are initially unindexed. Let us sequentially draw studies one-by-one, completely at random and label their indexes as 1,2,3,…M according to the order drawn. This labelling will not have any impact on the results (point estimate of effect size and its estimated variance) for any published method we know of.

 The classical random effects model for the true study effect size parameters for study j, Θj are assumed to satisfy:

 Θj = θ+ε j with E(ε j)=0. (2)

 This model is called “Effects-at-Random” because all of the true effect sizes in (2) are drawn with replacement from a single “urn” with mean θ.

 The primary goal is to estimate the overall effect size θ, based upon the effect size estimates for study j, defined by $\hat{θ}$j, which are presumed to be conditionally unbiased for Θj given the study identifier.

 The global estimator of θ is$ \hat{θ}$, defined by $\hat{θ}$=Σ Wj$\hat{θ}$ j with the weights Wj satisfying ΣWj=1.

 Good News: The random assignment of indexes (with or without the model) assures us that nonparametrically that all E($\hat{θ}$ j) are identical, due to the exchangeability of the M-pairs (Wj,$ \hat{θ}$j).

 Bad News: As we illustrated above, The Wj cannot be considered as constants, so the classical assumption that they are constants fails and this failure is well beyond any sampling errors due to estimating the so-called optimal weights.

 Note that this model does not accommodate the possibility that there is any association between study design parameters (such as weight or sample size) and effect size.

 We shall prove that only uniform weighting can **guarantee** that $\hat{θ} is$ unbiased. To see this, our labelling assures us that the M pairs (Wj $, \hat{θ}$ j) are exchangeable. This makes every $\hat{θ}$ j have the same mean θ and since the weights sum to 1.00, the same expected weight 1/M.. Let us study Cov(Wj $, \hat{θ}$ j).

 E($\hat{θ}$j) = θ and E(Wj)=1/M.

 Cov(Wj$, \hat{θ}$ j) = E(Wj$\hat{θ}$j)- E(Wj) E($\hat{θ}$j) = E(Wj $\hat{θ}$j)- θ/M.

 Hence,

 E(Σ Wj$\hat{θ}$j)= M Cov(Wj$, \hat{θ}$ j) +θ. (3)

 Thus, for $\hat{θ}$ to be unbiased, the **bias term** MCov(Wj$, \hat{θ}$j) =0 and this can only be guaranteed nonparametrically under equal weighting. Note that if we repeated the random label assignment, the bias term would not change. Also note that a positive (negative) covariance is associated with bias trends toward overestimating (underestimating) the mean response θ respectively.

 Next, note that since the mainstream uses equation (4) below to estimate the variance of the estimate $\hat{θ}$:

 Var( $\hat{θ}$)≈ ΣWj 2 Var($\hat{θ}$j) (4)

That this is **not a valid** **formula** , because it ignores the random nature of the weights, and this renders the inverse variance weights as optimizing an invalid metric for both fixed effects and effects-at-random.

The correct formula is:

 Var( $\hat{θ}$)≈ E[( ΣWj $\hat{θ}$j)2]-E2(ΣWj $\hat{θ}$j) (5)

which is quite different from (4).

 We conclude the following about the classical weighted estimates:

1. It is dangerous to presume there is no association (interaction) between the weights and the estimates.

(ii) The overwhelming majority of random-effects Meta-Analysis publications are suspect. Those with public health impact need to be looked at under more rigorous methodology.

 (iii) With this random labelling, as can be seen in our example, the conditional coefficient of variation of the weights (CCV), given the actual unindexed set of the M weights can be substantial and does not vanish when the number of studies, M, is large.:

 CCV=100 SQRT[M{ Σ {Wj $ $-(1/M)} 2 ] (6)

 **This major variability in the weights has not been recognized in the mainstream**. This conditional coefficient of variation can only be zero under uniform weighting.