Additional information provided to Committee members by EPA concerning the derivation of equation 4 on page 21 of the Agency's White Paper titled *Valuing mortality risk reductions for policy: a meta analytic approach*.

The w_{ij}'s are normalized to sum to one, so the estimator is a weighted average of the observations.[12] To derive optimal weights, we choose the w_{ij}'s to minimize the mean squared error of the estimator. First, we partition the error for each observation into group-specific and observation-specific components,

$$y_{ij} = y + \eta_i + \mu_{ij} + \varepsilon_{ij}, \tag{2}$$

where y is the true mean VSL among the general U.S. adult population (our target of estimation), η_i is a group-level non-sampling error, μ_{ij} is an observation-level non-sampling error, and ε_{ij} is an observation-level sampling error. We assumed that the expected value of each error component is zero and that all error components are uncorrelated.[13] The mean squared error of the estimator is

$$MSE = E[(\hat{y} - y)^2] = E\left[\left(\sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}(\eta_i + \mu_{ij} + \varepsilon_{ij})\right)^2\right]. \tag{3}$$

Taking the derivative of (3) with respect to each w_{ij}, setting the resulting expressions equal to zero, and imposing the constraint that the weights must add to one, we find that the optimal (MSE-minimizing) weights are

$$w_{ij} = \frac{\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right)^{-1}}{\sum_{i=1}^{I}\sum_{j=1}^{m_i}\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right)^{-1}}, \tag{4}$$

where σ_η^2 is the variance of the group-level non-sampling errors, σ_μ^2 is the variance of the observation-level non-sampling errors, and se_{ij}^2 is the sampling error variance for observation j from group i.

This section of hidden text shows how to derive equation (4) by minimizing MSE. First, write the mean squared error as:

$$MSE = E[(\hat{y} - y)^2] = E\left[\left(\sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}(y + \eta_i + \mu_{ij} + \varepsilon_{ij}) - y\right)^2\right].$$

Weights will sum to one, $\sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij} = 1$, so…

$$MSE = E[(\hat{y} - y)^2] = E\left[\left(\sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}(\eta_i + \mu_{ij} + \varepsilon_{ij})\right)^2\right].$$

[12] Constraining the weights to sum to one ensures that the estimator is consistent (assuming that the expected values of all error components are zero) but rules out shrinkage estimators, which in some cases can reduce the mean-squared error of the estimator by introducing some bias for a more-than-offsetting reduction in variance (e.g., Tibshirani 1996).

[13] Note that even though the error components are assumed to be uncorrelated, the total errors of the estimates within a group still will be correlated due to the common group-level error, η_i .

Next, assuming all error terms are uncorrelated we can write:

$$MSE = \sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}^2 E\left[\left(\eta_i^2 + \mu_{ij}^2 + \varepsilon_{ij}^2\right)\right]$$

Taking expectations, and assuming that the expected value of all errors is zero, we get:

$$MSE = \sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}^2\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right).$$

The first-order condition for an optimum is:

$$\frac{\partial MSE}{\partial w_{ij}} = 2w_{ij}\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right) - \lambda = 0,$$

where , λ is the Lagrange multiplier on the constraint that the weights must sum to one. This implies:

$$w_{ij} = \frac{\lambda}{2\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right)} \quad \forall ij.$$

And finally, using the fact that the weights sum to one to solve for λ and plugging the resulting expression back into the above equation, we get:

$$w_{ij} = \frac{\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right)^{-1}}{\sum_{i=1}^{I}\sum_{j=1}^{m_i}\left(m_i\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2\right)^{-1}}.$$

To estimate the se_{ij}'s, we used the standard error for each observation reported in their respective original studies (or calculated by us as described in Appendix B). The non-sampling error variances, σ_η^2 and σ_μ^2, are unknown, so the optimal weights in equation (4) cannot be applied in practice. However, below we describe estimators that we developed for the non-sampling error variances so that estimates of the optimal weights could be used.

In sections to follow we describe how we estimated σ_η^2 and σ_μ^2 using non-parametric and parametric approaches. In the remainder of this section we describe several alternative non-parametric estimators that can be derived as special cases of equation (4), which represent competing assumptions about the relative sizes of the variances of the error components associated with the observations. We used this set of alternative estimators to examine the robustness of the estimated VSL depending on the assumed nature of the relative precision of each observation and their possible correlations within groups, and to facilitate comparison to previous meta-analysis studies that may have used one of these alternatives as their primary estimators. The non-parametric estimates that we calculated are:

1. <u>Simple mean</u>: $\hat{y} = \frac{1}{N}\sum_{i=1}^{I}\sum_{j=1}^{m_i} y_{ij}$. This estimator would be optimal if the observation-level non-sampling error variance is much larger than the group-level non-sampling error