Value of Statistical Life Analysis and Environmental Policy: A White Paper

Chris Dockins, Kelly Maguire, Nathalie Simon, Melonie Sullivan
U.S. Environmental Protection Agency
National Center for Environmental Economics

April 21, 2004

For presentation to Science Advisory Board - Environmental Economics Advisory Committee

Correspondence:

Nathalie Simon
1200 Pennsylvania Ave., NW (1809T)
Washington, DC 20460
202-566-2347
simon.nathalie@epa.gov

Table of Contents

Value of Statistical Life Analysis and Environmental Policy: A White Paper

1 Introduction

The U.S. Environmental Protection Agency (EPA) uses a value of statistical life (VSL) estimate to express the benefits of mortality risk reductions in monetary terms for use in benefit cost analyses of its rules and regulations. EPA has used the same central default value (adjusted for inflation) in most of its primary analyses since 1999 when the Agency updated its *Guidelines for Preparing Economic Analyses* (USEPA, 2000). Prior to the release of the *Guidelines*, EPA sought advice from the Science Advisory Board's Environmental Economics Advisory Committee (SAB-EEAC) on the appropriateness of this estimate and its derivation. In 2000, EPA also consulted with the SAB-EEAC on the appropriateness of making adjustments to VSL estimates to capture risk and population characteristics associated with fatal cancer risks.[1] Currently, the Agency engaged with the SAB Advisory Council on Clean Air Act Compliance Analysis (the Council) on appropriate approaches to valuing mortality risks in the context of the 812 Second Prospective Analysis.[2]

EPA is now in the process of revising and updating its *Guidelines* and as such we are revisiting our approach to valuing mortality risk reductions. The literature has grown considerably since EPAs default estimate was derived and several EPA-funded reports have raised issues related to the robustness of estimates emerging from the mortality risk valuation literature. Furthermore, several meta-analyses have been conducted of this literature, providing new means of deriving central, default values for consideration. EPA's goal in bringing this issue to the SAB-EEAC is to seek expert opinion and guidance regarding the most appropriate way in which to proceed in updating the VSL estimate used to assess the mortality risk reductions from environmental policy.

It is important to note that this discussion focuses exclusively on mortality risk valuation. While we recognize the importance of morbidity and co-morbidity risks, the focus of this particular White Paper is on mortality; morbidity will be addressed at a future time.

To help inform the discussion, this paper provides background on current EPA practices for valuing mortality risk reductions, briefly summarizes the findings of three cooperative agreement reports on various segments of the literature, and reviews three recent meta-analyses that derive aggregate VSL estimates. The paper concludes with charge questions for consideration and discussion by the EEAC members. Full copies of the cooperative agreement reports and the meta-analyses are included in the Appendices.

2 Current Guidance and Practice for Valuing Mortality Risks

---

1 *An SAB Report on EPAs White Paper Valuing the Benefits of Fatal Cancer Risk Reductions*, #EPA-SAB-EEAC-00-013, July 27, 2000.

2 *Review of the Revised Analytical Plan for EPA's Second Prospective Analysis – Benefits and Costs of the Clean Air Act 1990-2020*, Draft Report, #EPA-SAB-COUNCIL-ACV-XXX-XX, March 5, 2004. Portions related to VSL are included as Appendix B.

Reductions in mortality risk constitute the largest quantifiable benefits category of many of EPA's rules and regulations. As such, mortality risk valuation estimates are an important input to most of the Agency's benefit-cost analyses.

EPA's *Guidelines* advise analysts to use a central VSL estimate of $4.8 million in 1990 dollars. Based on the gross domestic product (GDP) deflator this converts to approximately $6.2 million in 2002 dollars. This value is derived from 26 estimates assembled for EPA's first retrospective analysis of the Clean Air Act (USEPA, 1997). Each estimate is from a different study, with 21 of the estimates from hedonic wage studies and the remaining five derived from contingent valuation (CV) studies. The estimates range from $0.9 million to $20.9 million (2002 dollars) and the studies were published between 1976 and 1991. The estimates are fitted to a Weibull distribution that is often used in probabilistic assessments of uncertainty in EPA benefits calculations. Appendix A contains a list of the estimates used by the Agency and indicates the study from which each was derived.

Until 2003, the estimate from EPA's *Guideline*s was uniformly applied to mortality risk reductions across program offices. EPA recently used an estimate of $5.5 million (1999 dollars) in its analysis of reduced mortality from air regulations. The economic analysis for EPA's Proposed Inter-State Air Quality Rule describes the approach.

> The mean value of avoiding one statistical death is assumed to be $5.5 million in 1999 dollars. This represents a central value consistent with the range of values suggested by recent meta-analyses of the wage-risk VSL literature. The distribution of VSL is characterized by a confidence interval from $1 to $10 million, based on two meta-analyses of the wage-risk VSL literature. The $1 million lower confidence limit represents the lower end of the interquartile range from the Mrozek and Taylor (2000) meta-analysis. The $10 million upper confidence limit represents the upper end of the interquartile range from the Viscusi and Aldy (2003) meta-analysis.[3]

This approach has been considered by the Council as part of their review of the Analytic Plan for the second Clean Air Act Prospective Analysis. As noted above, the Council is currently drafting its final report on the Analytic Plan.

2.1 "Adjustments" to the Base VSL

While there are many risk and population characteristics that may affect VSL estimates, to date EPA makes few adjustments to base estimates. Based on advice from the SAB-EEAC [4] and other committees,[5] EPA analysts have adjusted the base VSL estimate to account for the effects

---

[3] *Benefits of the Proposed Inter-State Air Quality Rule*, EPA 452-03-001, January 2004.

[4] An SAB Report on EPA's White Paper *Valuing the Benefits of Fatal Cancer Risk Reduction*, EPA-SAB-EEAC-00-013, July 27, 2000.

[5] *Arsenic Rule Benefits Analysis: An SAB Review,* EPA-SAB-RSAC-01-008, August

of time. Specifically, future risk reductions valued according to VSL are discounted, including risk reductions spread over any latency period and/or cessation lag. This issue is of particular importance for cancer risks, but has also been employed for mortality from particulate matter.

Because income elasticity is believed to be positive, EPA has also adjusted current VSL estimates for anticipated income growth over time. Specific elasticity estimates have varied somewhat, but have been generally based on a review of the empirical literature on cross-sectional income elasticity of WTP. Income growth has been defined as the change in per capita GDP over time and projections of GDP growth are based on estimates from the Bureau of Labor Statistics.

EPA has been advised that the costs of illness for fatal cancers may be added to VSL estimates to assess the benefits of reducing cancer mortality.[6] The empirical effect of this addition is small and to date, the Agency has incorporated it only once into its regulatory analyses.

Finally, EPA has been advised that the evidence does not support empirical adjustments for other factors that may differ between study and policy cases, and that may affect VSL, including:

- risk preferences or risk aversion;
- age;
- cross-sectional income;
- cancer premium, fear, or dread;
- baseline health status; and
- voluntariness and controllability of risk.

2.2 Sensitivity and Alternate Estimates

The *Guidelines* allow for sensitivity analysis around key risk and population characteristics that affect the value of risk reduction. The particular parameters for a given sensitivity analysis should be guided by the benefit transfer concerns for that policy context.

EPA has considered several of the factors listed above in sensitivity analyses or alternative estimates. "Alternative estimate" is generally used to describe an analysis that incorporates scientific conclusions believed to be equally valid alternatives to the primary estimate. Sensitivity analyses typically employ other points on the Weibull distribution of VSL described in the *Guidelines*. For the case of the effect of age on VSL, EPA has employed various treatments including sensitivity analysis using the value of statistical life year, empirical adjustments based on CV studies, and an alternate analysis using only stated preference literature. The recent Durbin amendment to the appropriations bill for the Agency now precludes the Agency from performing any age-based adjustments when estimating the value of

---

2001.

[6] *Arsenic Rule Benefits Analysis: An SAB Review,* EPA-SAB-RSAC-01-008, August 2001 (p. 6).

mortality risk reductions to adults in most contexts.[7]

3 Robustness of Estimates from Mortality Risk Valuation Literature

In anticipation of periodically revisiting the Agency's approach to mortality risk valuation, EPA funded three studies to examine the various segments of the mortality risk valuation literature. Black et al. (2002) and Alberini (2004), provide empirical assessments of the robustness of mortality risk valuation estimates emerging from hedonic wage-risk studies and contingent valuation studies, respectively. Blomquist (2004) provides a summary of the averting behavior literature.[8] All three studies are provided in their entirety in Appendices B, C and D.

3.1 Hedonic Wage Literature

Black et al. (2002) systematically examines the robustness of hedonic wage estimates of willingness to pay for mortality risk reductions using data sets commonly used in this area of research. To perform an hedonic wage study researchers generally need information on worker characteristics, including wage, and job risk. Specifically, this study examines the roles of functional form, measurement error, and unobservable characteristics using various data sets, including data on occupational risk from the Bureau of Labor Statistics (BLS) and National Institute for Occupational Safety and Health (NIOSH) and data on worker characteristics from the Current Population Survey (CPS), Outgoing Rotation Groups of the CPS, and the National Longitudinal Survey of Youths (NLSY).

Since no large data set exists that contains both basic types of information, researchers must match observations from various sources, making decisions on how best to combine the data which are often reported at different levels of aggregation. For example, researchers can choose

[7] Public Law 108-199, "Consolidated Appropriations Act, 2004," Section 419 reads "None of the funds provided in this Act may be expended to apply, in a numerical estimate of the benefits of an agency action prepared pursuant to Executive Order No. 12866 or section 312 of the Clean Air Act (42 U.S.C. 7612), monetary values for adult premature mortality that differ based on the age of the adult."

[8] Blomquist 2004 appears in *Review of Economics and the Household* but is based on the work emerging from the cooperative agreement.

to create either industry-based or occupation-based risk measures to match with the worker-level data, each with its own difficulties. If industry-based measures are used, different occupations within an industry receive the same risk level (e.g., a miner and secretary for a mining firm). However, occupation-based measures potentially problematic because occupation is not well classified, with employers and employees often disagreeing on occupation classification.

3.1.1 Baseline estimates

The authors begin with ordinary least squares (OLS) estimation of simple log linear hedonic wage equations for three different worker samples and using both NIOSH and BLS risk data. The covariates included in the basic regression include basic controls such as worker age, education, union status, marital status, race and ethnicity. Also included, when possible, are variables to control for workers' firm size, state of residence, and one-digit industry and occupation. Results are reported separately for men and women. The positive VSL estimates that are calculated from these basic results range from $3.7 million to $16.4 million. The authors raise concerns regarding variation in other working conditions that may be captured in the estimates and interpret the instability they find in their parameter estimates as evidence that the measures of job risk are correlated with the regression error. The remainder of the paper is focused on identifying the source of this instability.

3.1.2 Role of Functional Form

The authors estimate the same equations using a more flexible functional form and using non-parametric approaches. In both cases they find that the results are just as volatile. Interestingly, they also find that the estimates are somewhat larger using the more flexible functional form. They conclude that the instability is not a result of the log linear specification. They also note that their tests do not necessarily mean that the non-linear specification is correct, only that it implies the presence of other problems.

3.1.3 Measurement Error

The authors note three possible sources of measurement error:

- Low sampling variation within industry and occupation cells given the small size of some of these cells (in recognition of this problem, BLS and NIOSH suppress data when number of fatalities is low);
- Heterogeneity in the actual job risk and non-random assignment of that job risk within occupation (e.g., late night convenience store clerks tend to be male and older);
- Industry and Occupation are not measured accurately, especially at three-digit level.

After using various techniques to determine the magnitude of the measurement error, they then attempt to correct or mitigate the error with limited success. Their efforts lead them to believe that the estimates they obtain are inconsistent and should not be used in policy analysis.

3.1.4 Unobservables

Using the National Longitudinal Survey of Youth (NLSY) data, the authors explore the effect of other characteristics not typically included in hedonic wage equations and typically not available in other worker samples, such as illegal drug use and Armed Forces Technical Qualification (AFTQ) scores. They find that those who admitted using illegal drugs tended to take on more occupational risk while those with higher AFTQ scores tended to sort into safer jobs. Hence, job risk is an endogenous variable.

3.1.5 Conclusions

In short, Black et al. find that results from hedonic applications to wage-risk data are not robust and are in fact quite unstable. For many of the specifications they try, they find a negative price of risk and for others they find that small changes in the covariates or risk measure used produce large variation in the estimated price. In their attempts to identify the source of this variation, they first examine the functional form of the regression equation. Using more flexible functional forms does not alleviate the problem. Second, they find "overwhelming evidence" that the job risk measures contain measurement error and that this error is correlated with covariates commonly used in the wage equations. Studies that do not correct for these errors would likely underestimate the value of risk reductions. Finally, they provide evidence that occupation risks are correlated with other characteristics typically not provided in the data sets commonly used for this type of analysis.

The findings of Black et al. are of obvious concern to EPA given the Agency's reliance to date on the hedonic wage-risk literature in determining its central, default VSL for use in policy analysis. To the extent that hedonic estimates are unstable, questions regarding the continued use of this literature in policy applications must be addressed.

<u>3.2 Contingent Valuation Literature</u>

Alberini (2004) examines the robustness of estimates of willingness to pay for mortality risk reductions derived from contingent valuation data and illustrates the empirical effects of some well-known problems in the contingent valuation literature. The author selects several papers from the literature and examines the robustness of the WTP estimates under alternative assumptions regarding (i) choice of distribution for WTP; (ii) presence of contaminating responses (yea-saying, nay-saying, and random responses); (iii) treatment of zero WTP; (iv) interpretations of WTP responses; (v) endogeneity of subjective baseline risks and/or risk reductions; (iv) treatment of regressors and outliers, and (vii) sample selection bias. Each issue is examined separately for some subset of the papers for which Alberini was able to obtain data.

The five CV studies from the original 26 studies in Viscusi (1991) are of obvious interest, but the author was able to obtain data for only one of the five. Additional studies are chosen from the relatively recent literature on the basis of quality, and Alberini's judgment of the study results' applicability to environmental policy, as well as availability of data. The studies used in Alberini (2004) are: Gerking, de Haan and Schulze (1988); Johannesson and Johansson (1996);

Johannesson, Johansson, and Lofgren (1997); Persson, Norinder, Hualte, and Gralen (2001); Krupnick, et al. (2002); and Alberini, Cropper, Krupnick and Simon (forthcoming).

3.2.1 Choice of Distribution

Analyzing single-bounded responses from two studies, Alberini finds that mean estimates may depend crucially on assumptions about the underlying distribution of responses, and on the coverage of the range of possible WTP. Median WTP is far less sensitive to these factors. Alberini concludes that double-bounded questions may be preferable and that median WTP should be used rather than means.

3.2.2 Mixture Models

"Mixture models" are presented to illustrate how one could model and estimate the extent of contaminating responses to a CV survey (e.g., 'yea-saying,' 'nay saying'). The models are estimated using data from three of the studies collected. The results are interesting, but it is clear that it is difficult to reliably estimate mixture models. Alberini concludes that contaminating responses could be an important factor affecting inferences of respondent behavior, and thus questionnaires should include debriefing questions designed to identify the presence of contaminating responses in such a way that the debriefing results can be used in the statistical analysis.

3.2.3 Treatment of Zero Responses

Alberini shows that alternative interpretations of zero responses can significantly affect the estimates of mean WTP, while again estimates of the median are not substantially affected. This issue is intimately related to the choice of the underlying distribution of responses, though it is confounded somewhat by the treatment of single- vs. double-bounded responses.

3.2.4 Treatment of Extreme Responses

Alberini also examines the effects of extreme responses on WTP estimates and shows that dropping outliers can have large effects on both the mean and median WTP, although median WTP is less sensitive.

3.2.5 Endogeneity of Risk

When surveys elicit perceptions of baseline risks or risk reductions, it is important to test and correct for endogeneity between subjective risks and WTP. If not controlled for, endogeneity biases estimates of the risk coefficient and confounds scope tests. Alberini shows that endogeneity can affect inferences regarding whether it is absolute or relative risk changes that determine WTP. This is important since values for absolute risk changes are needed in order to calculate a VSL. Alberini recommends, therefore, that researchers express risks in both absolute and relative terms. The author also examines the effect of excluding implausibly large subjective risk values and finds this can also affect the results.

3.2.6 Conclusions

Although Alberini (2004) does not provide a comprehensive examination of the contingent valuation literature, her findings are nevertheless of significant consequence to the Agency. Methods for eliciting willingness to pay values for mortality risk reductions have clearly advanced with time. Her systematic examination of a number of key issues using several available datasets in the analysis of CV data, as well as the presentation and interpretation of CV results, at the very least provides a number of factors that should be considered in selecting studies on which to base any central, default VSL estimate.

3.3 Averting Behavior Literature

Although not a formal meta-analysis or a detailed statistical treatment of the averting behavior literature, Blomquist (2004) summarizes the empirical averting behavior VSL literature and provides a heuristic review of existing estimates. The author begins by presenting a basic framework for estimating VSL based on averting behavior and follows with a brief review of existing estimates.

The study finds that VSL for adults from this literature ranges from a little less than $1.7 million to $7.2 million in 2000 dollars. Making a few assumptions, Blomquist finds a simple average adult VSL of approximately $4.5 million. In the author's judgment the range of "best estimates" is about $2 million to $7.2 million, with a subjective best estimate of $4 million. Blomquist reports evidence that VSLs may be greater, or at least not less, for children than for adults, but existing studies are not conclusive on this point. Furthermore, empirical evidence on VSLs for senior citizens is limited and not conclusive.

The author then makes some broad conclusions about the averting behavior literature:

- More recent estimates are larger than those in earlier studies. This is credited to greater use of hedonic approaches rather than relying on values of time, disutility costs, etc.;
- Difficulties with individual risk perception are an issue, but not a barrier for estimating VSL from averting behaviors. This conclusion is based on (i) evidence that individual risk perceptions are correlated with expert assessments, (ii) that VSL estimates can be adjusted for risk misperception in a sensitivity analysis (and these values may actually be preferred), and (iii) VSL estimates can be informative even if they are not "perfect.";
- VSLs from averting behaviors have tended to be somewhat lower than those from hedonic wage studies. However, the difference is not great, and there is reason to believe that hedonic wage VSLs are biased upward (e.g., Shogren and Stamland, 2002);
- VSLs from averting behaviors tend to be higher than those from stated preference studies. The paper attributes much of this difference to hypothetical bias in SP studies.;
- Blomquist suggests a meta-analysis of averting behavior VSL estimates, specifically recommending that the analysis consider: base risk level, risk change, adjustment for risk perception bias, value of time, treatment of disutility or jointness in consumption, and individual characteristics.

4 Meta-Analyses of the Mortality Risk Valuation Literature

Since EPA derived the VSL estimate cited in the *Guidelines* advances have been made in the field of mortality risk valuation. There are new examinations of how context affects mortality risk valuation, as well as new hedonic wage and contingent valuation studies. Some new CV studies make use of improved risk communication devices, which have been shown to improve the validity of these estimates. Key recent work on mortality valuation includes Krupnick, et al. (2002), Eeckhoudt and Hammitt (2001), Viscusi (2004), Smith, et al. (2003), and Smith, et al. (2004). While we recognize the important contributions these and other recent papers have made directly, we focus on three recent meta-analyses that include many relatively new mortality valuation studies.

Meta-analysis is a potentially useful means of combining individual but related studies in an analytically rigorous way that accounts for individual characteristics of each study. We reviewed a number of meta-analyses for this background paper. Each was assessed as to whether it provided a viable estimate or range of estimates of VSL that the Agency could use for policy analysis. The studies by de Blaij, et al. (2000), Miller (1990), and Miller (2000) lack the level of coverage and/or statistical rigor deemed appropriate for Agency use. We review three studies, however, in more detail, as they contain broad coverage of the available literature, rigorous statistical analyses, and/or a presentation of a range of predicted VSL estimates. These studies can provide useful insights into our efforts to update the VSL estimate used in EPA analyses. The three studies we review below are Kochi, Hubbell and Kramer (2003), Mrozek and Taylor (2002), and Viscusi and Aldy (2003).[9] Summaries of each of these studies appear below, including descriptions of the criteria used to select the individual studies used, data and statistical specifications, and results. Appendix J presents a combined bibliography of all the VSL studies included in the meta-analyses considered below.

4.1 Summary of Kochi, Hubbell and Kramer[10]

Kochi, Hubbell and Kramer (2003) employ an empirical Bayes estimation method to generate

---

[9] The Council in their assessment of the 812 Analytic Blueprint considered these same three studies for the Second Prospective Analysis.

[10] This summary is based on the 2003 version of the analysis that accompanied the EPA's Analytic Blueprint for Second Prospective Analysis. An updated version of the study is currently under review for publication and will be provided to the SAB-EEAC. We have not thoroughly assessed differences in the two versions.

predicted VSL estimates using multiple results from both hedonic wage and stated preference studies. To identify potential studies for inclusion, the authors searched for recent work in bibliographies from previously published meta-analyses and review articles, citations from other VSL studies, web searches, and personal contacts. They collected 47 hedonic wage studies and 29 contingent valuation studies for potential inclusion in their analysis.

In deciding whether to include a study, they applied the same criteria used in Viscusi (1992), a review article of 37 studies. Viscusi employed four explicit criteria for selecting studies in his analysis:

- include only hedonic wage and contingent valuation studies; consumer market studies "...failed to provide an unbiased estimate of the dollar side of the risk-dollar tradeoff, and tend to underestimate VSL." (p. 7);
- exclude hedonic wage studies using actuarial risk data (because these data include risks other than those on the job and therefore bias the VSL estimate down);
- include only studies using a simple regression estimation approach (as opposed to a more complex estimate of the tradeoff for discounted expected life years lost);
- studies must have a minimum sample size of 100.

In addition, the following selection criteria are noted as implicit in Viscusi (1992):

- only include hedonic wage studies for general or blue-collar workers;
- only include CV studies on samples of the general population;
- only include studies from high income countries (e.g., US, UK, Japan).

These selection criteria reduced the number of studies in Kochi, et al. from 76 to 45. They use all reported VSL estimates for reduced risk of immediate death from each study, resulting in 196 estimates.

Kochi et al. re-estimated all possible VSLs and associated standard errors for each included study based on information provided in each original study, using mean values for variables.[11] Recalculations that resulted in a negative VSL were excluded from the primary analysis, but are included in a sensitivity analysis.

The authors employed Bayes estimation, which requires the assumption that each estimate used be an independent sample. As this is unlikely if multiple observations from a single study are included, the authors array the culled VSL estimates into "homogeneous subsets" by author and other characteristics. A total of 60 subsets were created in this fashion, each assumed to be independent. Once subsets were created, a representative VSL for each subset was constructed by averaging VSLs and their standard errors within the subset. Predicted VSL estimates are based on these representative VSLs.

[11] The VSL from CV studies is calculated as WTP/(risk reduction).

This estimation method adjusts each of the representative VSLs based on within-study variability and the distribution of VSLs across studies. Smooth distributions are generated by using kernel density estimation, assuming a normal distribution for the kernel function. To test for sensitivity of the results to original valuation method, the authors separately estimate distributions for hedonic wage and contingent valuation studies. A bootstrap technique, resampling each sample of method-specific estimates 1000 times, is then applied to compare the different distributions of VSL.

4.1.1 Results

The primary results using all studies are summarized in Table 1. The table shows a mean VSL of $5.4 million with a standard error of $2.4 million (2000 dollars). A sensitivity analysis examining hedonic wage and CV estimates as separate sets found the hedonic wage distribution has a mean of $2.8 million (standard error = $1.3 million), while the hedonic wage distribution has a mean of $9.4 million (standard error = $4.7 million). The differences in means, medians, and interquartile ranges between the distributions are statistically significant. The sample containing U.S. studies only has a mean of $8.5 million (standard error = $4.9 million).

Table 1
Results of Empirical Bayes Estimates and Bootstrap Tests for Distribution Comparisons
Reproduced from Table 2 in Kochi, Hubbell, and Kramer (2003)

	Mean (million $)	Standard Error (million $)	Coefficient of Variation	Bootstrap Test		
				Mean	Median	Interquartile
Distribution Comparison by Evaluation Method						
Total (60)	5.4	2.4	0.4	P-value (Ho: HW = CV)		
CV (18)	2.8	1.3	0.5			
HW (42)	9.4	4.7	0.5	<0.001	<0.001	<0.008

4.1.2 Limitations

Although the study is useful for aggregate level comparisons, it does not account for the impact of specific study characteristics, including population characteristics, on VSL. Furthermore, the study gives no weight to the original authors'judgements to distinguish reasonable or preferred estimates from others, with the exception of negative VSLs. This may be statistically valid, but is troublesome because the conclusions of the authors who are most familiar with their research are lost.

It is not clear to what extent this analysis captures different specifications used across studies. If the VSLs are based on regressions with different specifications and this is not otherwise captured in the analysis, then it seems the "homogeneous groupings" are somewhat arbitrary and could be made differently. Since differences in specification are likely to significantly influence the resulting VSL estimates, the study should account for these differences in some way.

The authors also note that the results are sensitive to small VSLs with low variances. These estimates receive a great deal of weight in the empirical analysis. Removing Krupnick, et al. (2000), for example, increases the mean estimated VSL by almost 10%.

4.2 Summary of Mrozek and Taylor

Mrozek and Taylor's analysis focuses on results from hedonic wage studies only. Estimates from 47 studies were used although the authors do not specify how they selected their studies. Ultimately, 14 studies were subsequently dropped because:

- mean risk values, and in some cases also mean earnings, were not reported (6 studies);
- the risk measure confounds death and injury (1);
- observations represent industries, not individuals (1);
- study was unavailable (1);
- many variables were not reported (1);
- mean wages were incorrectly calculated (1);
- no VSL estimate was reported or obtainable (2 studies); or
- results were identical to another study (1).

As with the Kochi et al. study, multiple observations are used from each study when the original authors reported variations in model specifications or samples from which VSL estimates could be obtained. One to 28 observations are obtained from each study. Variables included in the meta-regressions are of three types: those which may influence wage/risk tradeoffs (e.g., mean hourly earnings, national unemployment rate in the year wage data was collected, mean annual risk of death); those describing the sample (e.g., if the data is from a national sample of US workers, if risk variable included a worker's self-assessment of risk, if the sample is 100 percent white collar); and methodological choices of the original researchers (e.g., if a risk-squared term is included, the number of industry categories controlled for, if at least one dummy variable describing a job characteristic was included).

The authors use weighted least squares so that each study, regardless of the number of observations drawn from it, is weighted equally. Four models are estimated, in each case the log of VSL is the dependent variable. Model 1 is the most inclusive, while model 2 eliminates observations based on samples with high risks and those using actuarial data. Model 3 further restricts the sample to the U.S. and includes a dummy variable indicating where five or more industries were controlled for in the original study. Model 4 is the same as model 3 except that it incorporates a continuous variable indicating the number of industries controlled for in the original study.

4.2.1 Results

All four models indicate a positive and significant relationship between the mean risk and VSL. The authors find this relationship to be concave - VSL estimates begin to decline when mean risk is between 1.2 to 1.67 deaths per 10,000, depending upon the model. The coefficient on earnings is positive but significant only in models 1 and 2. VSL estimates from national U.S. samples are higher than those from specialized U.S. samples and the use of NIOSH data results in higher VSL estimates than do estimates generated from BLS data.

The authors use the meta-analysis results to develop revised estimates of VSL by predicting VSL as if the original studies had all followed a set of best practice assumptions (e.g., including a risk-squared term, including at least one occupational dummy, including at least one dummy describing a job characteristic). Table 2 presents mean adjusted predictions from models 3 and 4 for five baseline risk levels ranging from 0.25 to 2 deaths per 10,000, by potential dataset (BLS or NIOSH), and by control for inter-industry differences. Estimates assuming the use of NIOSH data are higher than those assuming use of BLS data and range from $1.35 million to $11.7 million (1998 dollars), estimates decline for risks greater than 1.5 per 10,000. The authors conclude that the evidence best supports an estimate of $2 million at the average occupational risk level of 0.5 per 10,000. Refining this estimate for an average worker leads to an estimate of approximately $2.6 million (see footnote 17).

Table 2
Estimates of the Value of Statistical Life: Mean Adjusted Fitted Values[a]
Reproduced from Table 4 in Mrozek and Taylor (2002)

Risk (x 10^{-4})	Based on Model (3), Table 3		Based on Model (4), Table 3	
	< 5 Industries	**$ 5 Industries**	**0 Industries**	**7 Industries**
BLS Risk Data				
P = 0.25	$3.82m	$1.35m	$2.99m	$1.27m
	(1.39)	(0.47)	(1.12)	(0.40)
P = 0.5	$4.73m	$1.67m	$3.90m	$1.65m
	(1.64)	(0.53)	(1.44)	(0.51)
P = 1.0	$6.25m	$2.20m	$5.57m	$2.36m
	(2.36)	(0.73)	(2.22)	(0.80)
P = 1.5	$6.78m	$2.39m	$6.33m	$2.68m
	(3.02)	(0.92)	(2.83)	(1.03)
P = 2.0	$6.05m	$2.13m	$5.72m	$2.42m
	(3.09)	(0.92)	(2.83)	(1.03)
NIOSH Risk Data				
P = 0.25	$6.59m	$2.32m	$5.24m	$2.22m
	(2.62)	(1.00)	(2.08)	(0.84)
P = 0.5	$8.16m	$2.88m	$6.82m	$2.89m
	(3.17)	(1.20)	(2.72)	(1.10)
P = 1.0	$10.8m	$3.80m	$9.76m	$4.13m
	(4.57)	(1.65)	(4.18)	(1.68)

P = 1.5	$11.7m	$4.13m	$11.1m	$4.69m
	(5.65)	(1.95)	(5.21)	(2.07)
P = 2.0	$10.4m	$3.68m	$10.0m	$4.24m
	(5.57)	(1.85)	(5.06)	(1.97)

[a] Values are expressed in millions (1998 dollars). Standard errors are in parentheses.

4.2.2 Limitations

The study has been criticized in a paper by Harrison (2002) for failing to report standard errors and for the authors' choice of which estimates from each study they included. For example, the authors excluded estimates in original studies that were statistically insignificant or negative - such as the negative coefficients on the BLS variable in certain studies.

Hammitt (2002) and Krupnick (2002) each provide commentary on the Mrozek and Taylor study. Hammitt highlights several important findings. For example, Mrozek and Taylor find that failure to control for non-fatal risks is less significant than previous studies report and they confirm a common result that NIOSH data produces VSL estimates that are substantially higher than BLS data. Hammitt also highlights the importance of controlling for industry as a significant finding from Mrozek and Taylor. Hammitt notes the mixed evidence in Mrozek and Taylor concerning the use of actuarial versus perceived risk estimates, as well as the mixed results concerning pre- and post-tax dollars. Hammitt does question the Mrozek and Taylor results concerning the relationship between risk and VSL estimates. Specifically, Hammitt believes that the increase in VSL as risk increases is too large to be supported by standard models.

Krupnick (2002) focuses on the policy relevance of the Mrozek and Taylor meta-analysis. Mrozek and Taylor report a best estimate of $2 million, which is about 66 percent less than the estimate currently used by EPA in most benefit-cost analysis. While there are examples of rules that may have "failed" the benefit-cost test by using this lower estimate, Krupnick notes that there are many factors that enter into the decision-making process on a given policy, making it unlikely that this lower estimate would significantly change decision making in these cases. While Krupnick endorses the Mrozek and Taylor study, he does state that concerns with the hedonic wage literature may supplant the use of this study in policy analysis.

4.3 Summary of Viscusi and Aldy

Viscusi and Aldy conduct a review of more than 60 studies of mortality risk across 10 countries, examining a number of econometric issues, the effects of unionization on risk premiums, and the effects of age and income on VSL estimates. The analysis includes fifty-two hedonic wage studies from the U.S. and other countries selected based on the following set of criteria:

- written in English;
- published in academic journal or book;
- provides enough information to calculate a VSL.

The authors did not attempt to eliminate studies or modify the original VSL estimates. Point estimates from each study are those using the “whole sample” based on the original authors’ preferred specification.

The empirical analysis drops 3 studies that did not have an income measure. It also appears that three studies that did not report mean risks were dropped, resulting in 46 studies for OLS specifications. Other specifications dropped either one or two more studies, but it is not clear which ones. Values in the final set of studies range from $0.5 million to $20.8 million. Half of the U.S.-based studies estimate a VSL from $5 to $12 million. The median estimate from the sample is about $7 million.

In the statistical analysis, the authors first replicate four other published meta-analyses, using the preferred specifications of the authors of those studies (Liu, Hammit and Liu, 1997; Miller, 2000; Bowland and Beghin, 2001; Mrozek and Taylor, 2002).[12]

Next, the paper presents original meta-analyses employing six specifications, three using OLS specifications and three robust specifications with Huber weights. The simplest specifications include only the log of income and mean risk as dependent variables; two other specifications include mean risk squared; and the most complete and robust specifications also include variables to control for the underlying data source, whether risks are subjective, whether the study included a morbidity variable, and regional, urban, industry, and occupation dummies.

4.3.1 Results

The predicted values in the study are presented in Table 3, which is adapted from Table 8 of Viscusi and Aldy (2003). Generally, predicted values for the U.S. range from $5.5 million to $7.6 million. The study notes that median predicted values were generally very close to the means.

The authors predicted mean VSL estimates by using the estimated coefficients from the meta-analysis to predict the natural logarithm of VSL for each original study. After converting log(VSL) to VSL the study-specific predicted values were averaged to get the mean estimates presented in Table 3. Confidence intervals were constructed by using the prediction error for each study from the meta-analysis regressions. Lower and upper confidence intervals for each study were averaged to produce the lower and upper confidence intervals reported below. Predicted U.S. mean values are constructed based on regression samples using all countries, but with averaging across only U.S. studies. The authors note that the confidence intervals are valid only under the assumption that the model is specified correctly.

The meta-analysis is undertaken to estimate the effects of income on VSL and the study finds

[12] Liu, Hammit and Liu (1997) and Bowland and Beghin (2001) focus on developing countries and thus are not considered in our summary.

that income elasticity for VSL ranges from about 0.5 to 0.6 across several specifications. The authors note that the 95 percent confidence interval on income elasticity never exceeded 1.0.

Table 3
Mean predicted VSL, U.S. sample
Reproduced from Viscusi and Aldy (2003)

	OLS 1	OLS 2	OLS 3	Robust w/ Huber wts	Robust w/ Huber wts	Robust w/ Huber wts
Variables	Log(Y) & mean risk	OLS1 + mean risk^2	Full set	Log(Y) & mean risk	OLS1 + mean risk^2	Full set
Mean predicted VSL (95% C.I.)	5.5 (3.8 - 8.1)	5.8 (4.1 - 8.3)	6.9 (3.1 - 16.2)	6.1 (4.6 - 8.2)	6.3 (4.8 - 8.4)	7.6 (3.0 - 19.4)

4.3.2 Limitations

While the meta-analysis results are highly consistent across specifications, the confidence intervals for the regressions that include the full set of covariates are broad because there are relatively few degrees of freedom. Moreover, the precise VSL values used for each study in the sample are not fully clear. The paper reports VSL's for each study in the analysis, but some of these are in the form of a range. Finally, the selection criteria does not include estimates from "grey" or unpublished literature.

5 Conclusion and Summary

Since 1999 EPA has relied on a central VSL estimate of $6.2 million (2002 dollars) for most of its economic analyses, which is derived from a Weibull distribution of 26 hedonic wage and contingent valuation studies of mortality risk valuation. Recently, in air regulations EPA has used an estimate of $5.5 million (2003 dollars), which is derived from recent meta-analyses. In light of additions and advances in the literature, the time is ripe for revisiting the VSL estimate(s) used in EPA policy analysis.

This background paper reports on three cooperative agreements that assess the hedonic wage, contingent valuation, and averting behavior literatures, as well as reviews three recent meta analyses of the mortality risk valuation literature.

Each of the cooperative agreements highlights areas of concern with the particular literature under investigation. Black, et al. (2003) raise concerns with the stability of hedonic wage estimates, given the large changes in results that come from slight changes in specification or choice of data. Alberini (2004) demonstrates how modifications in specification can affect results and asserts that median estimates are more stable than mean estimates, though researchers must be attuned to the impact of outliers and zero values when doing their estimation. Finally, Blomquist (2004) reviews the averting behavior literature and encourages a more thorough analysis for use in policy decisions.

We also review three recent meta-analyses of the mortality risk valuation literature. Kochi, Hubbel and Kramer (2003) use Bayesian techniques to combine contingent valuation and hedonic wage studies in a meta-analytic framework. They recalculate the original estimates to account for independence and report an estimate of $5.4 million from their studies. Both Mrozek and Taylor (2003) and Viscusi and Aldy (2004) conduct meta analyses of the hedonic wage literature. The studies differ in their selection criteria and how they use the estimates. Mrozek and Taylor report a best estimate of $2 million, while Viscusi and Aldy report a best estimate of around $6 million.

These reports and studies are informative as EPA revisits the best VSL estimate to use in policy analysis. This background paper concludes with a series of Charge Questions to guide discussion of issues confronted when using the existing mortality risk valuation literature to evaluate environmental policies.

References

Alberini, Anna. 2004. Willingness to Pay for Mortality Risk Reductions: A Re-examination of the Literature. *Draft Final Report,* Cooperative Agreement with U.S. Environmental Protection Agency, #015-29528.

Alberini, Anna; Maureen Cropper, Alan Krupnick, and Nathalie Simon. forthcoming. Does the Value of a Statistical Life Vary with Age and Health Status? Evidence from the United States and Canada. *Journal of Environmental Economics and Management*.

Black, Dan A., Jose Galdo and Liqun Liu (2003). *How Robust Are Hedonic Wage Estimates of the Price of Risk?* Final Report to the USEPA [R 829-43-001].

Black, Dan A. and Thomas J. Kneiser. 2003. On the Measurement of Job Risk in Hedonic Wage Models. *Journal of Risk and Uncertainty* 27(3): 205-220.

Blomquist, Glenn C. 2004. Self-Protection and Averting Behavior, Values of Statistical Lives, and Benefit Cost Analysis of Environmental Policy. *Review of Economics of the Household* 2: 89-100.

Bowland, B.J. and J.C. Beghin. 2001. Robust Estimates of Value of a Statistical Life for Developing Economics. *Journal of Policy Modeling* 23: 385-396.

de Blaij, Arianne, Raymond J.G.M Florax, Piet Rietveld, and Erik Verhoef. 2000. The Value of Statistical Life in Road Safety: A Meta-Analysis. *Tinbergen Institute Discussion Paper*, TI 2000-089/3.

Eeckhoudt, Louis R. and James K. Hammitt. 2001. Background Risks and the Value of a Statistical Life. *Journal of Risk and Uncertainty* 23(3): 261-279.

Gerking, Shelby; Menno de Haan and William Schulze. 1988. The Marginal Value of Job Safety: A Contingent Valuation Study. *Journal of Risk and Uncertainty* 1(2): 185-200.

Hammitt, James K. 2002. Commentary On: What Determines the Value of Life?: A Meta-Analysis. *Journal of Policy Analysis and Management* 21(2): 271-73.

Harrison, Glenn W. 2002. The Use and Abuse of Meta-Analysis in Governmental Policy. *University of Central Florida Working Paper.*

Johannesson, Magnus and Per-Olov Johansson. 1996. To Be or Not to Be, That is the Question: An Empirical Study of the WTP for an Increased Life Expectancy at an Advanced Age. *Journal of Risk and Uncertainty* 13: 163-174.

Johannesson, Magnus; Per-Olov Johansson and Karl-Gustav Lofgren. 1997. On the Value of Changes in Life Expectancy: Blips Versus Parametric Changes. *Journal of Risk and Uncertainty* 15: 221-239.

Kochi, Ikuho; Bryan Hubbell and Randall Kramer. 2003. An Empirical Bayes Approach to Combining Estimates of the Value of a Statistical Life for Environmental Policy Analysis. *Unpublished Manuscript*.

Krupnick, Alan. 2002. Commentary On: What Determines the Value of Life?: A Meta-Analysis. *Journal of Policy Analysis and Management* 21(2): 275-282.

Krupnick, Alan K.; Anna Alberini, Maureen Cropper, Nathalie Simon, Bernie O'Brien, Ron Goeree, and Martin Heintzelman. 2000. Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents. Washington, D.C.: *Resources for the Future Discussion Paper 00-37*, Resources for the Future.

Krupnick, Alan; Anna Alberini, Maureen Cropper, Nathalie Simon, Bernie O'Brien, Ron Goeree, and Martin Heintzelman. 2002. Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents. *Journal of Risk and Uncertainty* 24: 161-184.

Li, J-T.; James K. Hammitt, and J-L. Liu. 1997. Estimating Hedonic Wage Function and Value of Life in a Developing Country. *Economics Letters* 57: 263-275.

Miller, Ted R. 1990. The Plausible Range for the Value of Life - Red Herrings Among the Mackerel. *Journal of Forensic Economics* 3(3): 17-39.

Miller, Ted R. 2000. Variations Between Countries in Values of Statistical Life. *Journal of Transport Economics and Policy* 34(2): 169-188.

Mrozek, Janusz R. and Laura O. Taylor. 2002. What Determines the Value of Life? A Meta-Analysis. *Journal of Policy Analysis and Management* 21(2): 253-270.

Persson, Ulf; Anna Norinder, Krister Halte, and Katarina Gralen. 2001. The Value of a Statistical Life in Transport: Findings from a New Contingent Valuation Study in Sweden. *Journal of Risk and Uncertainty* 23(2): 121-134.

Shogren, Jason F. and Tommy Stamland. 2002. Skill and the Value of Life. *Journal of Political Economy*, 110(5): 1168-73.

Smith, V. Kerry; Mary F. Evans, Huyn Kim, and Donald H. Taylor, Jr. 2004. Do the Near-Elderly Value Mortality Risks Differently? *The Review of Economics and Statistics*, 86(1).

Smith, V. Kerry; Subhrendhu K. Pattanayak, and George L. Van Houtven. 2003. VSL Reconsidered: What do Labor Supply Estimates Reveal About Risk Preferences? *Economics Letters*, 80(2): 147-153.

U.S. Environmental Protection Agency (USEPA). 2000. *Guidelines for Preparing Economic Analyses*. Office of the Administrator, EPA 240-R-00-003, September.

U.S. Environmental Protection Agency (USEPA). 1997. *The Benefits and Costs of the Clean Air Act: 1970 to 1990*. Office of Administration and Resources Management and Office of Policy, Planning and Evaluation, October.

Viscusi, W.K. 2004. The Value of Life: Estimates with Risks by Occupation and Industry, *Economic Inquiry*, 42(1): 29-48.

Viscusi, W.K. 1992. *Fatal Tradeoffs: Public and Private Responsibilities for Risk.* New York, NY: Oxford University Press.

Viscusi, W. Kip and Joseph E. Aldy. 2003. The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World. *Journal of Risk and Uncertainty* 27(1): 5-76.

Charge Questions for Discussion

The charge questions are structured around a set of broad issues that define the general objectives of this review.

I. Literature support for a revision of the current Guidelines for valuing changes in fatal risk.

In 1999, the Science Advisory Board - Environmental Economics Advisory Committee reviewed the draft *Guidelines for Preparing Economic Analyses*. The *Guidelines* state that the Agency would continue to conduct periodic reviews of the risk valuation literature and revise the Guidelines accordingly, under advisement from the SAB. Though the literature has grown since the publication of the 2000 *Guidelines*, the Agency's practice of valuing changes in fatal risks has largely been unchanged. Does the literature support a revision of the current Guidelines for valuing fatal risk changes?

II. Questions on the important strengths and limitations of the available literature and how these factors be accounted for in practice.

A. The Background Paper summarizes several EPA commissioned reports that document methodological concerns underlying VSL studies that use hedonic wage equations, contingent valuation surveys, and averting behavior methods. What are the important practical lessons EPA can draw from these reports, and how should these be used to evaluate literature to be used by EPA?

B. To what extent is it scientifically appropriate for the Agency to use VSL estimates from unpublished studies and studies from developing countries in developing mortality risk valuation policy?

III. The risk valuation literature has grown substantially since the 1999 Guidelines were published. EPA has questions about what is the most scientifically appropriate way for EPA to aggregate the literature in updating its mortality risk valuation policy. There are a number of alternatives to consider:

A. Current Practice (fitting a distribution)

EPA *Guidelines* recommend using a distribution of VSL estimates based on 26 studies from the literature. A Weibull distribution was fit to the set of estimates, yielding a central estimate of approximately $6.1 million. Is this sort of "curve fitting" a preferred methodology for deriving a distribution of VSL values for use in economic analyses of EPA regulations?

B. Meta-analyses

(i) There are three widely-circulated meta-analyses of VSL estimates that are recent contributions to the literature. Is meta-analysis the preferred methodology for deriving VSL values for use in economic analyses of EPA regulations?

(ii) The white paper summarizes three widely-circulated meta-analyses of VSL estimates that appear to be generally regarded as high quality.[1] These analyses differ in their selection criteria, the scope of studies they consider, and their technical approach to combining existing VSL estimates. In general, what are the relative strengths and weaknesses of each study in regards to application to EPA policy analyses? Does one of these studies emerge as a preferred candidate for VSL estimates for EPA policy analyses?

(iii) Each of the three studies use different criteria to select estimates to include in the analysis (e.g., only HW studies, HW and CV studies). Are there particular selection criteria that should be required in any meta-analysis used by EPA for policy analysis?

(iv) Similarly, each of the studies uses different statistical techniques to calculate their VSL estimates. For example, some studies rely on regression techniques, whereas others fit a particular distribution to the data. What approach should EPA use for calculating VSL estimates for policy analysis?

(v) Each of the meta-analyses manipulates the original data to some extent. For example, some studies adjust for after-tax wages, whereas others do not. Is there a set of such manipulations that the EEAC believes to be critical for any meta-analysis? Are there some data manipulations that are generally incompatible with sound meta-analysis?

(vi) How should a quality meta-analysis handle zero or negative VSL estimates from studies that otherwise meet its selection criteria for inclusion?

(vii) If the Agency relies upon multiple meta-analyses to estimate VSL for policy analysis, how can the different meta-analyses most rigorously and appropriately be combined given that they use different statistical procedures, and overlapping, but not identical sets of studies?

C. Are there other alternatives methodologies EPA should consider for aggregate the literature in updating its mortality risk valuation policy?

---

[1] The three studies are Viscusi and Aldy (2003), Kochi, Hubbell and Kramer (2003), and Mrozek and Taylor (2000).

IV. The characteristics of risks and populations addressed in the VSL literature are often different from those addressed by EPA policies. The SAB has addressed some of these questions concluding that the only empirically feasible adjustments to a base VSL are (1) discounting over periods of latency and cessation lag, and (2) increasing VSL over time to account for rising real income.

A. Does the literature continue to support empirically accounting for these effects in policy analysis?

B. Does the literature support empirically accounting for other risk and population characteristics in transferring existing VSL estimates to the analysis of EPA policies?

V. Empirical analysis is always limited by data constraints. The analysis by Black, et al., for example, highlights the impact of existing data limitations in hedonic wage studies. EPA is interested in hearing from the SAB-EEAC members on how the Agency might assist research through efforts to make data more available.

A. Can useful analytical gains be made through low-cost improvements in data quality or increased data availability? What steps can EPA and other government agencies take in the short term to facilitate research through improved data quality or increased accessibility to existing data sets?

B. The EEAC recently reviewed EPA's draft Environmental Economics Research Strategy and provided advice regarding research needs for mortality valuation as part of that review. Given the additional information provided to the committee for this review, do EEAC members wish to identify any additional research needs or provide any modifications to their recent advice?

C. What do members of the EEAC see as the most fruitful, long-term strategies for overcoming the challenges of using the existing literature for environmental policy analysis?