EPA-SAB-16-xxx

The Honorable Gina McCarthy
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: SAB Review of EPA's Proposed Methodology for Updating Mortality Risk Valuation Estimates for Policy Analysis

Dear Administrator McCarthy:

The EPA's National Center for Environmental Economics (NCEE) requested advice from the Science Advisory Board (SAB) on proposed improvements in the agency's methodology for estimating the value of mortality risk reductions, also known as the value of statistical life (VSL). The EPA requested that the SAB review three documents: (1) a white paper titled *Valuing Mortality Risk for Policy: a Meta-Analytic Approach* ("White Paper"); (2) a report titled *The Effect of Income on the Value of Mortality and Morbidity Risk Reductions*; and (3) a technical memorandum titled *Recommended Income Elasticity and Income Growth Estimates: Technical Memorandum.* The White Paper was developed to describe the EPA's proposed approach for estimating values for reductions in mortality risk for use in benefit-cost analysis. The other documents discuss options for updating the agency's recommended estimate for the income elasticity of the VSL.

In response to the EPA's request, the SAB Environmental Economics Advisory Committee was convened to review the White Paper and other documents. The SAB was asked to respond to 17~~9~~ charge questions organized into six topics focusing~~that focused~~ on: (1) whether the methods used to select data for the analysis were appropriate and scientifically sound; (2) whether relevant studies were adequately included in the analysis; (3) whether the methodology used to analyze the data was scientifically sound; (4) whether the EPA's VSL estimates represented scientifically sound ~~conclusions~~; (5) the development of a protocol for future updates of the VSL; and (6) whether the EPA's approach for estimating the income elasticity of VSL was appropriate and scientifically sound. The enclosed report provides the SAB's consensus advice and recommendations.

In general, the SAB finds that the meta-analytic methods used in the White Paper to develop VSL estimates for the U.S. population appear to be ~~scientifically sound and~~ broadly consistent with standard and accepted practices but the SAB recommends improvement and clarification of the methods in several areas. ~~However,~~ As explained in the enclosed report, more detailed information should be provided in the White Paper to explain and justify use of the methods, discuss how standard and accepted practices have been applied, ~~and~~ modify and strengthen the analysis, and enable readers to replicate calculations. ~~The SAB also recommends that the agency consider using a variation of the non-~~

Commented [K1]: Several people at the end of today's call mentioned the importance of consensus to the perception of the VSL estimates; My concern is that we are being asked based on the standards **today for defensible estimates** –how does this body of evidence fare?

1.Consequentiality is now central to the credibility of SP estimates; it is never mentioned in the white paper
2. The literature post Chay and Greenstone (JPE, 2005) has focused on the information used for identification

EPA cannot help that the existing literature for the most part did not effectively deal with either issue; the white paper makes no mention of either issue –I checked.

The VSL is the single most important number EPA uses –at some point it will experience significant scrutiny. I do not want EPA's fine record in benefit cost analyses to be tainted by a failure to recognize the limitations in the information they have available or by incomplete documentation of the assumptions used to re-organize old estimates to fit new circumstances. When these method extend beyond conventional practice and there is no independent evaluation of them, we need to say that. When old methods for organizing past research are recognized as flawed –it seems to me we need to raise the questions with them.

Commented [K2]: We would have trouble answering this comepletely and would need to enumerate that fact at the outset

Commented [K3]: Again there is not sufficient information to answer this questions details are missing; we would have to say that at the outset

Commented [K4]: We would have to question this and again not give a complete answer –so on 3 of the 4 boad themes we are not answering fully what is asked and we would have to say that

Commented [K5]: No they are not

Commented [K6]: More than clarification

~~parametric method that incorporates information on sampling error variance from each study to develop estimates of the VSL.~~ The SAB's major comments and recommendations are provided below.

- The evidence of study validity considered by the EPA in developing the dataset for the analysis is ~~appropriate but~~ incomplete. To strengthen the assessment, the EPA should consider applying additional tests of validity. The EPA should ~~and~~ also clarify how evidence of validity was applied to all of the studies considered for use in the analysis.

- The SAB recommends that in the future, the EPA broaden the scope of studies used to derive values for reducing both mortality and morbidity risks. There are a significant number of published studies that estimate willingness to pay for improved health and reduced health risks, and a literature on benefit-risk and risk-risk tradeoff preferences in health and health care and reduced risk for highway fatalities that could enrich the evidence on risk preferences and provide support for benefits-transfer applications.

- There has been a lack of significant growth in the number of studies used by the EPA to estimate the VSL ~~literature~~ since the last consideration of this topic by the SAB in 2011. The SAB provides citations for additional studies that could be included in the White Paper. However, the SAB also recommends that the agency consider commissioning more studies or creating other incentives for new studies to improve the prospect for a deeper literature to support future reviews of VSL.

- Some VSL estimates in the White Paper were constructed by weighting subpopulation estimates to approximate an estimate for the general population. Given the limited VSL literature, the SAB recognizes the need to develop a weighting approach for subpopulation estimates. However, additional information is needed in the White Paper to explain how the weighting was actually done and how the studies were brought together for the aggregate estimate~~. Moreover, the~~As a result, the White Paper mixes discussion of two kinds of procedures, population weighting and benefit transfer. EPA's analysis should be modified to ensure that population weighting is accomplished using standard procedures and that benefit transfer assumptions and procedures are appropriately described and applied.

- The White Paper should provide detailed information about how the standard error of the VSL is calculated when the original studies do not report it. A detailed description of the method, including the formula used in the calculation for each study, should be provided in the White Paper.

- The White Paper classifies estimates into independent samples called groups. The SAB supports grouping the studies in the White Paper based on similar samples to account for the lack of independence in estimates constructed from the samples. However, additional detail should be provided to clarify how the grouping decisions were made and an analysis should be conducted to check the robustness of the results to different plausible group definitions.

- The SAB finds that a five year interval for updating VSL estimates is appropriate, but there is a need to increase the pool of high quality studies to support the VSL meta-analysis. The EPA should consider whether estimation of VSL and its various attributes should be a high priority topic for grants and fellowships, sponsored conferences, and special issues of journals. There should be a

Commented [K7]: There is no literature cited that describes the basis for developing weighted means as a benefit transfer practice. This point is different from the process of using survey weights to allow a population mean to be constructed from a sample.

Commented [K8]: This will be affected by adjustments for income and other estimates that are "attached to empirical tradeoff measures

Commented [K9]: It may be reasonable to consider establishing a criterion for when a sufficient body of new research is available to warrant an update. In the absence of peer reviewed research describing new procedures for organizing "old" studies, one argument would hold that analysis time to consider new strategies is not warranted and the only adjust to the established VSL would be for the cost of living.

single set of criteria for determining which studies are of sufficient quality to be included in current and future estimates of VSL.

The EPA report and technical memorandum on the income elasticity of VSL provide reasonable summaries of the income elasticity literature. However, the summary of the literature indicates that there is not an adequate informational basis for deriving a consensus estimate of the income elasticity of VSL. Furthermore, there has been relatively little growth in median income over the last two decades, particularly for groups represented in the samples used for hedonic wage studies. Therefore, adjustment of VSL estimates by an income elasticity of VSL and index of income growth (based on GDP per capita) ~~does not seem to be~~ is not appropriate. Further justification for this strategy based on peer reviewed research would be warranted to proceed with this logic. However, conversion of VSL to inflation adjusted dollars is appropriate. ~~Therefore the SAB recommends that the EPA consider using the preferred VSL model specification to compare VSL estimates at different points in time and use that to obtain the implied income elasticity of VSL.~~

Commented [K10]: And perhaps declines in real income for the groups whose risk tradeoffs have been measured

The SAB appreciates the opportunity to provide the EPA with advice on this important subject. We look forward to receiving the agency's response.

Sincerely,

NOTICE

This report has been written as part of the activities of the EPA Science Advisory Board (SAB), a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The SAB is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names of commercial products constitute a recommendation for use. Reports of the SAB are posted on the EPA Web site at http://www.epa.gov/sab.

U.S. Environmental Protection Agency
Science Advisory Board
Environmental Economics Advisory Committee

CHAIR
Dr. Madhu Khanna, ACES Distinguished Professor in Environmental Economics, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, IL

MEMBERS
Dr. Kevin Boyle, Professor and Director, Department of Agricultural and Applied Economics, Program in Real Estate, Virginia Polytechnic Institute and State University, Blacksburg, VA

Dr. Sylvia Brandt, Associate Professor, Department of Resource Economics, University of Massachusetts, Amherst, MA

Dr. Richard Carson, Professor, Economics, Department of Economics, University of California, San Diego, La Jolla, CA

Dr. J.R. DeShazo*, Associate Professor for Public Policy, School of Public Policy and Social Research, University of California at Los Angeles., Los Angeles, CA

Dr. Mary Evans, Associate Professor, Robert Day School of Economics and Finance, Claremont McKenna College, Claremont, CA

Dr. Wayne Gray, Professor, Department of Economics, Clark University, Worcester, MA

Dr. Timothy Haab*, Department Chair and Professor, Department of Agricultural, Environmental and Development Economics, Ohio State University, Columbus, OH

Dr. F. Reed Johnson, Senior Research Scholar, Center for Medical and Genetic Economics, Duke Clinical Research Institute, Duke University, Durham, NC

Dr. Matthew Kotchen, Associate Professor, School of Forestry and Environmental Studies, Yale University, New Haven, CT

Dr. Matthew Neidell, Associate Professor, Department of Health Policy and Management, Mailman School of Public Health, Columbia University, New York, NY

Dr. James Opaluch, Professor and Chair, Department of Environmental and Natural Resource Economics, College of the Environment and Life Sciences, University of Rhode Island, Kingston, RI

Dr. Daniel Phaneuf, Associate Professor of Agricultural and Applied Economics, Department of Agricultural and Applied Economics, University of Wisconsin-Madison, Madison, WI

*Did not participate in the review of the EPA's proposed methodology for updating mortality risk valuation estimates for policy analysis.

Dr. Andrew Plantinga, Professor, Bren School of Environmental Science and Management University of California, Santa Barbara, Santa Barbara, CA

Dr. Richard Ready, Professor, Department of Agricultural Economics and Economics, Montana State University, Bozeman, MT

Dr. V. Kerry Smith, Emeritus Regents' Professor and Emeritus University Professor of Economics, Department of Economics, W.P Carey School of Business, Arizona State University, Tempe, AZ

Dr. Stephen Swallow, Professor, Department of Agricultural and Resource Economics, University of Connecticut, Storrs, CT

Dr. George Van Houtven, Director, Ecosystem Services Research, ~~Research Triangle Institute~~RTI International, Research Triangle Park, NC

Dr. JunJie Wu, Emery N. Castle Professor of Resource and Rural Economics, Department of Agricultural and Resource Economics, Oregon State University, Corvallis, OR

Dr. Jinhua Zhao*, Professor, Department of Economics, Department of Agricultural, Food and Resource Economics, Michigan State University, East Lansing, MI

SCIENCE ADVISORY BOARD STAFF
Dr. Thomas Armitage, Designated Federal Officer, Science Advisory Board, U.S. Environmental Protection Agency, Washington, DC

*Did not participate in the review of the EPA's proposed methodology for updating mortality risk valuation estimates for policy analysis.

Table of Contents

Acronyms and Abbreviations v

1. EXECUTIVE SUMMARY 1

2. INTRODUCTION 10

3. RESPONSES TO EPA'S CHARGE QUESTIONS 12

3.1. Meta-Analysis Dataset 12
3.1.1. Evidence of Validity for the Stated Preference Studies 12
3.1.2. Construct of the Risk Variable in Hedonic Wage Studies 20
3.1.3. Estimates of Value of Immediate Risk Reduction 23
3.1.4. Empirical Studies 25
3.1.5. Population Weighting in EPA's Analysis 27
3.1.6. Estimation of Standard Errors 36

3.2. White Paper Analysis 36
3.2.1. Overall Methodology for Analyzing the Data 36
3.2.2. Grouping Samples for Analysis 38
3.2.3. Addressing Sampling and Non-Sampling Errors 39
3.2.4. Non-parametric and Parametric Approaches for Estimating Value of Statistical Life 40

3.3. White Paper Results 42
3.3.1. Proposed Estimates of Value of Statistical Life 42
3.3.2. Influence Analysis 43

3.4. Protocol for Future Revisions of Value of Statistical Life 44
3.4.1 Criteria for Inclusion and Exclusion of VSL Estimates in Future Analyses 44
3.4.2. Valuing Reductions in Risks of Cancer 50

3.5. Income Elasticity of the Value of Statistical Life 52
3.5.1. Income Elasticity Literature 53
3.5.2. Analysis of Very Low Income Elasticity Estimates 53
3.5.3. Study Selection Criteria and Alternative Approaches for Estimating Central Income Elasticity of Value of Statistical Life 54
3.5.4. Income Elasticity of the Value of Non-fatal Health Effects 58

REFERENCES 59

APPENDIX A: THE EPA'S CHARGE QUESTIONS A-1

APPENDIX B: BIBLIOGRAPHY ON WILLINGNESS TO PAY IN HEALTH AND HEALTH CARE B-1

APPENDIX C: BIBLIOGRAPHY ON BENEFIT-RISK AND RISK-RISK TRADEOFF PREFERENCES IN HEALTH AND HEALTH CARE C-1

Acronyms and Abbreviations

BLS	U.S. Bureau of Labor Statistics
CFOI	Census of Fatal Occupational Injuries (U.S. Bureau of Labor Statistics)
~~COFI~~	~~Cost of Funds Index~~
CPS	Current Population Survey
CV	Contingent Valuation
FDA	U.S. Food and Drug Administration
FES	Fixed Effect Size
GDP	Gross Domestic Product
IEVSL	Income Elasticity of Value of Statistical Life
NOAA	National Oceanic and Atmospheric Administration
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RES	Random Effect Size
STAR	Science to Achieve Results Program
VRR	Value of Risk Reduction for Mortality
WTA	Willingness to Accept
WTP	Willingness to Pay

1. EXECUTIVE SUMMARY

The National Center for Environmental Economics in the EPA Office of Policy requested advice from the SAB on proposed improvements in the agency's methodology for estimating benefits associated with reduced risk of mortality. This methodology ~~takes into account~~ estimates the dollar amount that individuals are willing to pay for small reductions in mortality risk. The resulting values are combined into an estimate known as the value of statistical life (VSL) which is used in regulatory benefit-cost analysis. The EPA also requested that the SAB review options for accounting for changes in VSL over time as income grows, known as income elasticity of the VSL. The EPA submitted three documents to the SAB for review: (1) a white paper titled *Valuing Mortality Risk for Policy: a Meta-analytic Approach* (hereafter referred to as the "White Paper"); (2) a report by Robinson and Hammitt (2015) prepared for the EPA Office of Air and Radiation titled *The Effect of Income on the Value of Mortality and Morbidity Risk Reductions*; and (3) an EPA memorandum titled *Recommended Income Elasticity and Income Growth Estimates: Technical Memorandum.* The White Paper was developed to describe the EPA's proposed approach for estimating values for reductions in mortality risk for use in benefit-cost analysis. This approach includes assembling a VSL dataset from the published stated preference and hedonic wage study literature and using non-parametric and parametric analytic methods to develop central estimates of the average VSL among the general U.S. adult population. The EPA report and technical memorandum ~~on income elasticity of VSL~~ discuss options for updating the agency's recommended estimate for the income elasticity of the VSL.

The EPA asked the SAB to review the White Paper and other documents and respond to ~~19~~17 charge questions organized into six topics focusing ~~that focused~~ on: (1) whether the methods used to select the data set for the analysis were appropriate and scientifically sound, (2) whether relevant empirical studies were adequately captured in the White Paper, (3) whether the methodology used in the White Paper to analyze the data represents an appropriate and scientifically sound application of meta-analytic methods to derive VSL estimates; (4) whether the EPA's proposed VSL estimates represent reasonable and scientifically sound conclusions; (5) development of a protocol for future updates of the VSL; and (6) whether EPA's approach to estimating the income elasticity of VSL was appropriate and scientifically sound. This executive summary highlights the findings and recommendations of the SAB in response to the charge questions provided in Appendix A.

Commented [K11]: Cannot answer fully

Commented [K12]: Hicksian vs Marshallian issue

Commented [K13]: Cannot answer fully

Commented [K14]: Cannot answer fully

Commented [K15]: Cannot answer fullky especially issue of then applying them with GDP percaptia

Question 1a - Evidence of Validity of the Stated Preference Studies

The SAB was asked to comment on whether the methods EPA used in the White Paper to assess the validity of studies and value estimates were appropriate and scientifically sound. The SAB finds that the evidence of study validity considered by the EPA is ~~appropriate but~~ incomplete. To strengthen the assessment of study validity, the agency ~~should consider a broader set of criteria for validity. In addition, the EPA~~ should clarify how their criteria of validity were applied to all of the studies that were considered for use in the analysis. To better inform a weight of evidence decision to include or exclude a study, the EPA should expand the consideration of evidence of validity to include answers to additional key questions. This should include consideration of whether a valuation scenario was consequential. EPA should also fully document the validity evidence considered and how this evidence was used to include each study or value estimate. In addition, all future updates of the VSL should consider whether conditions for investigating study validity should be updated. For example, at the time of the last VSL review, consequentiality was not a predominant feature for evaluating study validity and it was not

standard practice to investigate the sensitivity of meta-analysis estimates to study and value estimate inclusion/exclusion in the analysis. These are important, new features of the VSL update and there are likely to be other new innovations in the literature at the time of the next update of the VSL.

Question 1b - Construction of the Risk Variable in Hedonic Wage Studies

The SAB was asked to comment on whether the hedonic wage studies included in the White Paper constructed the risk variable in a manner appropriate for use in the meta-analysis. In the White Paper the EPA used hedonic wage studies and estimates where the risk measure is differentiated by industry and at least one other characteristic (e.g., occupation, gender, age). The SAB supports excluding from the analysis those studies that employ fatality risk measures based on industry category alone. However, the current inclusion criterion that restricts the analysis to studies based on risk measures differentiated by industry and at least one other characteristic is inappropriate. Differentiating an industry level risk measure by some additional characteristics, for example age and gender, may lead to wage-risk tradeoff estimates unequally influenced by wage discrimination. In the short–run, the SAB recommends that: (1) the EPA alter its inclusion criteria to restrict its analysis to hedonic wage studies that employ fatality risk measures differentiated by occupation; and (2) the EPA include in the white paper a summary of recent meta-analyses of hedonic wage studies. This summary could provide information on the likely sensitivity of the final VSL measure to variations in the set of studies included in the calculations without having to replicate the research efforts already completed. In the long-run, the SAB recommends the EPA compile, make publicly available (e.g., on an internet webpage), and regularly update data that would encourage future revealed preference VSL research. To this end, the SAB recommends that the EPA pursue a Memorandum of Understanding with the U.S. Bureau of Labor Statistics to access and gain permission to make publicly available appropriate data that consistently merge the records from the Current Population Survey (CPS) with the Census of Fatal Occupational Injuries (CFOI). The SAB also recommends that in the long-run, the EPA apply a consistent hedonic wage model using data from the CPS and CFOI to generate comparable annual measures of VSL. Comparisons among these annual estimates could yield an estimate of the income elasticity of VSL. The SAB also recommends that the EPA pursue research to examine the various biases associated with hedonic wage studies, including an assessment of the adjustment needed to convert the willingness to accept measure obtained from hedonic wage studies to a Hicksian willingness to pay measure. ~~finds that the use of "industry and one other characteristic" risk measures is not appropriate for characteristics such as gender or age. This is because the risks are likely to be too varied within an industry group and because wage discrimination might affect the wage-risk differential across some groups. The SAB recommends that the analysis to determine the VSL only include hedonic studies where the risk variable includes variation by occupation, either with respect to occupation and industry or with respect to occupation only.~~

Commented [K16]: These data would involve the CPS data conventionally used to estimate hedonic wage models merged in a variety of ways with fatality risk measures distinguished by industry and occupation. These could include the risk measures for previous years as well as for the year of the wage survey.

Commented [K17]: Insights into the modeling and assumptions required to derive an estimate of the income elasticity of the VSL from these estimates of equilibrium tradeoffs. Need to re-consider in light of new focus on identification

Question 1c - Estimates of Value of Immediate Risk Reduction

The SAB was asked to comment on whether appropriate estimates from the stated preference literature were used in the White Paper to estimate the marginal willingness to pay for reduced risk of immediate death. The SAB has provided citations for several additional studies that could be included in the White Paper. In addition, the SAB finds that the supplementary analysis in one of the studies the EPA selected for use, Viscusi, Huber, and Bell (2014), does not provide clear evidence of ~~study validity (i.e.,~~ sensitivity of scope~~)~~. Moreover, the SAB recommends that the EPA broaden the scope of studies used to

derive values for reducing both mortality and morbidity risks. There are a significant number of published studies that estimate willingness to pay for improved health and reduced health risks, ~~and~~ a literature on benefit-risk and risk-risk tradeoff preferences in health and health care, and transportation literature on reduced risk for highway fatalities that could enrich the evidence on risk preferences and provide support for benefits-transfer applications. ~~The SAB also finds that discounting does not correctly account for the effect of time on VSL. The EPA should use a more correct construct such as the value of statistical life-years lost rather than the present value of a future statistical death.~~The SAB also finds that simple discounting does not account for confounded morbidity values in converting future deaths to equivalent immediate death values. The EPA should account for morbidity values in converting future mortality risks to equivalent instantaneous risks

Question 2 - Empirical Studies

The SAB was asked to comment on whether relevant empirical studies in the stated preference and hedonic wage literatures are adequately captured in the White Paper. There has been a lack of significant growth in the number of studies used by the EPA to estimate the VSL ~~literature~~ since the last consideration of this topic by the SAB in 2011. The SAB recommends that EPA search more broadly for additional studies not restricted to hedonic or stated preference methods. This could include an evaluation of whether studies using experimental or quasi-experimental methods may offer insight to VSL. Citations for several studies are provided. While these studies differ in methodology, data, or approach from studies already included in the White Paper, they offer potentially valid insight to the estimation of VSL. ~~As previously indicated, the SAB has provided citations for several additional studies that could be included in the White Paper. However, the SAB also recommends that t~~The EPA ~~agency consider~~ may need to commission~~ing~~ more studies or create~~ing~~ other incentives for new studies to improve the prospect for a deeper literature to support future reviews of VSL.

~~The White Paper should also contain more detail or information to allow readers to assess how the reliance on published studies, particularly other meta-analyses (including studies that drew from international data), might lead to results that differ due to publication bias, lags in publication, or other concerns.~~

Question 3 - Population Weighting in EPA's Analysis

The SAB was asked to comment on whether the population-weighting approach used in the White Paper to approximate a VSL estimate for the general population is appropriate and scientifically sound. Some estimates in the meta-analysis dataset in the White Paper are constructed by weighting subpopulation-specific estimates within a study in order to approximate an estimate for the general population. Given the limited VSL literature, the SAB recognizes the need to develop a weighting approach to use subpopulation estimates of VSL in the analysis. However, additional information is needed in the White Paper to explain in detail precisely how the weighting was actually done and how the studies were brought together for the aggregate estimate. The SAB recommends that EPA provide information sufficient to: (1) allow a third party to replicate the approach; and (2) distinguish between the use of population weights to derive a representative estimate of the VSL observation drawn from a particular study and the strategy of transferring benefit estimates from a source study as input to an estimate of VSL for some population (or timeframe) not directly addressed by the source study. Appendix B of the White Paper mixes discussion of two kinds of procedures, population weighting and benefit transfer. In

Commented [K18]: I would go further these are benefit transfer methods not population weighting

Commented [K19]: In some cases it appears that the studies involved used nonlinear properties (interaction terms) in the equilibrium wage models to estimate specific VSL tradeoffs for subgroups. As a result the estimates are all from the same model. They do not parallel the logic used in population weighting where sampling weights are specifically designed to reflect the sampling process in relationship to the population to be represented. In other cases, primarily the stated preference studies separate estimates are developed for sub-samples but the weights applied appear to be based on the fractions of these sub-population in a general census and not the sampling weights. Indeed the timing of the census may not correspond to the actual dates when the estimates for the VSL measures were derived. For these reasons it appears these are benefit transfer assumptions not the development of population mean measures based on survey weights. It is important to identify the peer reviewed research that describes the properties of these methods.

some cases, e.g., contingent valuation studies, the weighting procedure used by the EPA is comparable to using approximations for sampling weights. However, the procedure used for the hedonic wage studies is a benefit transfer. It is important that population weighting be accomplished using standard procedures and that benefit transfer assumptions and procedures for implementation be described and distinguished. The White Paper should also contain: (1) a more detailed explanation of how weighting procedures would affect estimates of standard errors; (2)~~. In addition: (1~~) an explanation of the implications of selection bias in survey-based studies that could result in exclusion of some members of the intended population~~should be explained;~~ ; ~~(2) greater consideration should be given to details of the specific studies being weighted;~~ and (3) connection of weights ~~should be connected~~ to the time periods of the original studies; ~~(4) there should be an adjustment for income differences in the populations in individual studies; (5) the EPA should explain how Hicksian and Marshallian measures of VSL were aggregated; and (5) the weighting process used for the specific studies listed in Appendix B of the White Paper should be clarified.~~ It is also important that income adjustments of the VSL estimates derived from hedonic wage and contingent valuation studies be consistent with the income concept relevant to each model. For hedonic wage models income is endogenous. However, for ~~contingent valuation this~~stated preference studies, this is not the case. ~~and expected utility is held constant.~~ The analysis of proper treatment of income should reconcile the~~se~~ modeling assumptions (including the role of income in Hicksian or Marshallian-based analyses) used by source studies with the use of any approach to deriving an estimate of VSL based on that study, before applying ~~some~~any adjustment. The EPA should also consider undertaking future work to investigate the possibility of developing a more complex set of subpopulation weights that build upon what is known about the subpopulations covered in each of the available studies. ~~***<<Chair's note: please clarify this sentence>>***~~

Commented [K20]: Needs to evaluate the logic underlying this approach ot ebenefits transfer

Question 4 - Estimation of Standard Errors

In the White Paper, the EPA attempts to estimate the standard errors of the VSL when the original studies do not report them. The SAB was asked to comment on whether the methods used to estimate these standard errors are appropriate and scientifically sound. There are two major issues that should be addressed with regard to estimation of standard errors. The first issue involves calculation of the standard error of the VSL when the standard error is not reported in the original study. The second, issue concerns the methods used to estimate standard errors for the overall VSL estimates in the White Paper. The SAB finds that the white paper does not provide sufficiently detailed information about how the standard error of the VSL is calculated when the original studies do not report it. EPA should document precisely how the standard error of the VSL is estimated when the original study does not report one, so that an independent party could replicate the calculations. ~~A detailed description of the method, including the formula used in the calculation for each study, should be provided in the White Paper.~~ The White Paper ~~also~~ uses a bootstrap approach to estimate standard errors for overall non-parametric VSL estimates. The SAB proposes an alternative, ~~perhaps~~ theoretically better, approach ~~way~~ to calculating~~e~~ standard errors for both the ~~each~~ non-parametric and parametric VSL estimator.

Commented [K21]: The issues raised with the construction of weights, attaching estimates in one year to populations in another building in adjustments for income growth all affect the computation of the standard erros; we discussed this I did not feel the qualifications were adequately reflected in this summary judgment

Question 5 - Overall Methodology for Analyzing the Data

The SAB was asked to comment on whether the methodology used in the White Paper to analyze the data represents an appropriate and scientifically sound application of meta-analytic methods to derive generally applicable VSL estimates for environmental policy analysis. In general, the SAB finds that the

meta-analytic methods used in the White Paper appear to be ~~scientifically sound and broadly consistent~~ with standard and accepted practices for conducting meta-analyses. ~~However, the SAB recommends that the White Paper more explicitly discuss the standard and accepted practices for conducting meta-analysis and how they have been applied. In particular, the EPA should: (1) refer to standard "fixed effect size (FES)" and "random effect size (RES)" methods, that are fundamentally similar to the "sampling error" and "total error" variance weighted mean approaches used in the White Paper, and describe how the White pPaper departs from these more standard practices; (2)~~The SAB recommends that the White paper provide more detail about each of the primary studies and the selected value estimates that would allow an independent party to replicate the results and that ~~in a way that~~ reinforces the direct comparability of the objects/commodities being valued; (3) discuss and make adjustments for differences in value estimates or~~in the~~ effect size measures across studies ~~of value estimates~~ (e.g., ~~the types of assumptions that would be needed to convert~~ hedonic studies provide Marshallian measures and stated preference studies provide ~~to~~ Hicksian measures); (4) conduct non-parametric and parametric analyses without a direct income adjustment~~ing~~ to VSL (there is insufficient evidence in the income elasticity of VSL literature to adjust VSL values from different studies to account for differences in income; furthermore, gross domestic product per capita, which was used as the measure of income, has not increased for some income groups) ~~values to account for differences in income but include an income measure as an explanatory variable in the parametric meta-regression~~; and (5) justify the use of "sample size weighted mean" in the non-parametric analysis to account for heterogeneity in the variance of effect size estimates.

Commented [K22]: Are not consistent with established practices. Adjustment of outcome measures (effect size) to assure that they are consistent across studies in consistent with conventional practice. However the scope of the adjustment undertaken in the white paper extends significantly beyond any practices that this committee was familiar with in the literature. As a result we could not conclude that they are consistent with peer reviewed

Question 6 - Grouping Samples for Analysis

The White Paper classifies estimates into independent samples, also called groups. Estimates from some hedonic wage studies that use the same or very similar worker samples are grouped together for the analysis. Similarly, some of the stated preference estimates using the same sample are grouped together. The SAB was asked to comment on whether this methodology represents an appropriate and scientifically sound approach for accounting for potential correlation of results that rely on the same underlying data. The SAB supports grouping the studies in the White Paper based on similar samples to account for the lack of independence in estimates constructed from the samples. However, additional detail should be provided to clarify how the grouping decisions were made. The SAB also recommends that the EPA conduct additional analysis to check the robustness of the results to different plausible group definitions. This robustness check should include: (1) exploring the sensitivity of results to alternative group assignments (e.g., grouping studies that used the same data set or the econometric approach together); (2) using the influence analysis to examine the robustness of results to excluding each group; and (3) identifying the primary estimate from each study and re-estimating the meta-regression using only primary estimates.

Question 7 - Addressing Sampling and Non-Sampling Errors

The White Paper presents an expression that characterizes optimal weights that account for sampling and non-sampling errors. The SAB was asked to comment on whether this is an appropriate and scientifically sound approach for addressing sampling and non-sampling errors. Additional information is needed to fully address this question. Derivation of the expression characterizing optimal weights that account for sampling and non-sampling errors should be explained in a more transparent way in the White Paper. Therefore, the SAB recommends including the various steps required to derive equation 4

in Section 4.1. The report should include the precise equation that is used by the EPA and citations that establish ~~Citations establishing~~ the validity of the basic approach.~~,~~ ~~if not the literal equation, should also be included~~. With regard to use of the weights, clarification of and justifications for the assumptions regarding the error components should be included in the White Paper. In addition, the SAB recommends that, without compromising best known, science-based practice for quantitative estimation, transparency be applied as a criterion for selecting an estimator.

Question 8 - Non-parametric and Parametric Approaches for Estimating Value of Statistical Life

The White Paper adopts both non-parametric and parametric approaches to estimate a VSL. The SAB was asked to comment on whether these approaches span a reasonable range of appropriate, scientifically sound, and defensible approaches to estimating a broadly applicable VSL for environmental policy and whether there are other methods that are more appropriate than those used in the White Paper. The SAB finds that additional information is needed in the White Paper to explain how these approaches were applied, particularly the nonparametric approach. Calculations should be documented with sufficient detail to allow a reader to know precisely how to replicate the calculations. Citations should be provided for the non-parametric approaches and better justification should be provided to explain why the methods used are relevant to finding the central tendency of VSL estimates from studies that, in most cases, report multiple estimates. The SAB supports the EPA's conclusion that the mean of group means estimator is the preferred non-parametric method because it has the smallest estimated standard error. It is important, ~~H~~however, ~~the EPA should also~~ to emphasize the rationale that ~~justify~~ use of the mean of group means estimator ~~on the grounds that it~~ avoids giving too much~~more~~ weight to studies~~papers~~ that report multiple~~more~~ estimates. In addition, the SAB recommends that: (1) the EPA explore the use of an alternative non-parametric method that incorporates information on sampling error variance from each study; (2) for the parametric estimator, the EPA provide a better explanation of and justification for the included control variables; and (3) the EPA should not include a time trend variable in either the parametric or nonparametric models, but should consider a sensitivity analysis to determine whether older or newer studies have a strong influence on the average VSL~~be consistent in its treatment of the time trend in VSL estimates~~.

Commented [K23]: The time trend may simply be reflecting the effects of the adjustments made which modify the effect size based on judgements associated with growth in income over time.

Question 9 - Proposed Estimates of Value of Statistical Life

Commented [K24]: Question 8 and 9 are estimated without really knowing how the input data to the process are computed

The White Paper presents VSL estimates using parametric and non-parametric models, pooled across stated preference and hedonic wage studies as well as balanced (i.e., giving equal weight to each study type), and weighted using different approaches. The EPA has proposed using the non-parametric model balanced mean of study means VSL estimate and the parametric model balanced VSL estimate. The SAB was asked to comment on whether these proposed estimates represent reasonable and scientifically sound conclusions from the analyses in the White Paper and whether there is a different set (or sets) of results that are preferable based on the data and analysis in the White Paper. As previously stated, the EPA's VSL estimates for the U.S. population were developed using a meta-analytic approach that appears to be consistent with standard and accepted practice. However, as indicated in the response to Charge Question 8, the SAB recommends that the EPA also explore the use of an alternative ~~consider using the~~ non-parametric method that incorporates information on sampling error variance from each study~~sampling error variance weighted group mean in place of the non-parametric mean of group means estimator~~. The SAB also recommends that the documentation of income adjustment to VSL be clarified in the White Paper. Adjustment of VSL estimates by an income elasticity of VSL and index of income

Commented [K25]: NO

Commented [K26]: Cannot be accepted as given

growth (based on GDP per capita) does not seem to be appropriate. However, conversion of VSL to inflation adjusted dollars is ~~would be~~ appropriate.

Question 10 - Influence Analysis

The results section of the White Paper concludes with an influence analysis. The SAB was asked to comment on whether the influence analysis is a reasonable way to characterize the influence of individual studies on the estimated VSLs, whether the results of the influence analysis suggest any changes or modifications to the EPA's estimation approach, and whether it is important to include an influence analysis. The SAB agrees that some form of influence analysis is important for meta-analysis in cases where there are few studies to consider, and therefore one or two individual studies might have a substantial influence on the estimates. Influence analysis of the maximum likelihood stated preference estimates in the White Paper indicates that the Corso, Hammitt and Graham (2001) study is well over two times more influential than the second most influential study. Therefore, the SAB recommends that the EPA consider using a robust estimation technique that limits the influence of this observation. The SAB also recommends that the EPA consider the potential for using regression diagnostic indexes for the parametric modeling of VSL.

Commented [K27]: Influence analysis assumes and error structure that is different than the one for meta data –only and informal gauge of potentially influential observations

Question 11 - Criteria for Inclusion and Exclusion of VSL Estimates in Future Analyses

The SAB was asked to comment on relevant statistical criteria for the inclusion of additional eligible VSL estimates and/or the exclusion of older VSL estimates that could help inform the development of a standardized protocol for future updates and the timing or frequency of those updates. The SAB finds that a five year interval for updating VSL estimates is appropriate, but there is a need to increase the pool of high quality studies to support the VSL meta-analysis. To accomplish this, the EPA should: (1) consider whether estimation of VSL and its various attributes should be a high priority topic for EPA grants and fellowships, sponsored conferences, special issues of journals, and awards; (2) consider the feasibility of sponsoring a refereed journal that focuses on analyses of direct relevance to meeting the agency's needs; and (3) obtain more general information about protocols for updating estimates from the experience of other agencies that construct economic index numbers for policy.

The SAB also recommends that: (1) there should be a single set of criteria for determining which studies are of sufficient quality to be included in current and future estimates of VSL; (2) the EPA should not restrict studies used for updating VSL to those published in peer-reviewed journals (studies outside of the peer-reviewed journals should be considered for inclusion following a transparent and rigorous peer review process) ***<<Chair's note: this statement should be discussed by the Committee>>*** ; (3) the EPA should consider whether useful information can be extracted from other studies that could improve estimates of VSL and its characteristics (e.g., latency, morbidity); (4) the EPA should not exclude studies based on non-national samples from use in updating VSL as long as: they~~re is a set of studies~~ are part of ~~that as~~ a group that is representative of the nation as a whole or they can be used to either develop a representative estimate for the nation as a whole or improve the representation of VSL values of subpopulations that are underrepresented or omitted from studies used to estimate a representative value for the nation as a whole ; and (5) the EPA should consider a long term strategy of requiring that a more inclusive set of research results, and even whole data sets, be made generally available for use by the research community and by government agencies.

Question 12 - Valuing Reductions in Risks of Cancer

The SAB was asked to comment on whether the selection criteria for identifying studies for valuing reductions in risks of cancer mortality should differ from those used in the current White Paper. The SAB was also asked whether the literature supports a non-zero differential between valuation of cancer and other mortality risk. Based on available studies, the SAB concludes that there is not sufficient evidence at this time to justify a non-zero cancer differential. EPA should encourage and support ongoing research on whether willingness to pay to reduce the risk of an early death preceded by a period of morbidity is correctly valued by summing the value of the morbidity plus the value of the mortality. At this time, the SAB does not have evidence to suggest that approach would over- or under-state the true willingness to pay. Nevertheless, ~~T~~the SAB recommends that, instead of adopting a non-zero cancer differential, the EPA consider using existing methods to value the morbidity that occurs prior to an early death and add that estimated morbidity value to conventional estimates of the value of the associated mortality. The EPA currently values morbidity from cancer in cases where the cancer is not fatal, but does not value morbidity in fatal cancer cases. The EPA should value cancer morbidity regardless of whether that morbidity leads to an early death. This recommendation also applies to other environment-related mortality risks, including cardio-pulmonary disease. ~~In addition, the EPA should encourage and support ongoing research on whether willingness to pay to reduce the risk of an early death preceded by a period of morbidity is correctly valued by summing the value of the morbidity plus the value of the mortality. At this time, the SAB does not have evidence to suggest that approach would over- or under-state the true willingness to pay.~~

Question 13 - Income Elasticity Literature

The SAB was asked to comment on whether the report by Robinson and Hammitt (2015) and the EPA Technical Memorandum provide an appropriate and scientifically sound summary of the income elasticity of VSL and income elasticity of non-fatal health effects literatures. The SAB finds that the Robinson and Hammitt (2015) and the EPA document *Technical Memorandum: Income Elasticity* provide reasonable summaries of the income elasticity literature. However, the SAB has provided citations for some additional studies that should be included in the summary of the literature. Very few studies have been conducted on the income elasticity of the value of statistical life. Going forward, the SAB recommends that the EPA support research to provide methodological guidance~~The SAB therefore recommends that the EPA consider the~~ that may enable use of~~possibility of using~~ estimates of the income elasticity for other related goods and services (such as consumer products that can be used to reduce health risks, and various forms of health insurance) to infer estimates of the income elasticity of the value of statistical life. ~~Going forward, the EPA should support more research to provide methodological guidance and empirical estimates in this important area.~~

Commented [K28]: It does not deal with the fundamental question of the ability to identify estimates of the income response basesd on equilibrium estimates for a marginal tradeoff and not structural parameters

Question 14 - Analysis of Very Low Income Elasticity Estimates

The "balanced" approach in the EPA Technical Memorandum on estimating income elasticity of VSL does not include reported mean estimates of zero, but does include very low reported mean estimates. The SAB was asked to comment on: (1) whether this was an appropriate and scientifically sound choice, and (2) how very low, non-zero, mean reported income elasticity results should be addressed in the EPA's analysis. The SAB finds that it is highly implausible for the income elasticity of VSL to be zero or negative. However, it is not completely clear how such estimates should be addressed in the EPA's

Commented [K29]: No basis for dropping estimates

analysis. It can be argued that such estimates may be theoretically impossible and therefore should be dropped from the analysis, but it is hard to find statistical justification for dropping them. One statistical justification for dropping them, however, is that the income in these studies was not well measured, which may bias the estimates toward zero. Imprecision in the quality of measurement will be partly reflected in the standard errors of the individual income elasticity of statistical life estimates. ~~The SAB finds that it is highly unlikely for the income elasticity of VSL to be zero or negative. However, to address the issue of low/zero estimates, the SAB recommends that, i~~Instead of calculating an unweighted mean of income elasticity of VSL estimates, the EPA could ~~should~~ use standard errors of individual income elasticity of VSL estimates to calculate a ~~weighted mean. The EPA should also calculate the~~ weighted mean of the income elasticity of VSL both with and without the low/zero estimates to assess their influence.

Commented [K30]: In the absence of addressing the identification issue this question cannot be resolved.No justification for dropping them

Questions 15 and 16 - Study Selection Criteria and Alternative Approaches for Estimating Central Income Elasticity of Value of Statistical Life

The SAB was asked to comment on whether the study selection criteria applied in the paper by Robinson and Hammitt (2015) are appropriate and scientifically sound and whether the additional inclusion of Viscusi, Huber, and Bell (2014) in the EPA Technical Memorandum is appropriate based on results reported in the study's on-line appendix. In addition, the SAB was asked to comment on two proposed alternatives for arriving at a central income elasticity of VSL. Robinson and Hammitt (2015) have done an admirable job summarizing the available literature. However, their analysis~~,~~ indicates that there is not an adequate informational basis for deriving a consensus estimate of the income elasticity of VSL. The inclusion of Viscusi, Huber and Bell (2014) does not alter this conclusion. The SAB finds that neither of the two alternatives put forward in Robinson and Hammitt (2015) and described in EPA's technical memorandum represent an adequate basis for providing an estimate of the income elasticity of VSL for policy purposes. Therefore the SAB recommends that the EPA consider the alternative approach of using the preferred VSL model specification to obtain and compare VSL estimates at different points in time and use that to obtain the implied income elasticity of VSL.

Question 17 - Income Elasticity of the Value of Non-fatal Health Effects

The EPA's Technical Memorandum recommends using the income elasticity of VSL to estimate income elasticity for the value of non-fatal health risks. The SAB was asked to comment on whether this represents an appropriate and scientifically sound approach given the available data. The SAB does not ~~fully~~ support using the income elasticity of VSL to estimate income elasticity for the value of non-fatal health risks because it is conceptually incorrect to apply income elasticity for one good to some other good. The SAB recommends that the EPA ~~explore~~ consider using~~e of~~ the income elasticity of expenditures on private health care products as a better proxy for the income elasticity of non-fatal health risks.

Commented [K31]: Near clear statement on need for theoretical links between averting behavior or other approaches to use income elasticites estimated in one context for another

2. INTRODUCTION

The National Center for Environmental Economics in the EPA Office of Policy requested advice from the SAB on proposed improvements in the agency's methodology for estimating benefits associated with reduced risk of mortality. This methodology ~~takes~~ estimates ~~into account~~ the amounts that individuals are willing to pay for reductions in mortality risk. The resulting values are combined into an estimate known as the value of statistical life (VSL) which is used in regulatory benefit-cost analysis. The EPA also requested that the SAB review options for accounting for changes in VSL over time as income grows, known as income elasticity of the VSL. The EPA submitted three documents to the SAB for review: (1) a white paper titled *Valuing Mortality Risk for Policy: a Meta-analytic Approach* (hereafter referred to as the "White Paper"); (2) a report by Robinson and Hammitt (2015) prepared for the EPA Office of Air and Radiation titled *The Effect of Income on the Value of Mortality and Morbidity Risk Reductions*; and (3) an EPA memorandum titled *Recommended Income Elasticity and Income Growth Estimates: Technical Memorandum.*

The White Paper was developed to describe the EPA's proposed approach for estimating values for reductions in mortality risk for use in benefit-cost analysis. This approach includes assembling a VSL dataset from the published stated preference and hedonic wage study literatures, and using non-parametric and parametric analytic methods to develop central estimates of the average VSL among the general U.S. adult population. The EPA report and technical memorandum on income elasticity of VSL discuss options for updating the agency's recommended estimate for the income elasticity of the VSL.

The White Paper provides context for the documents submitted to the SAB for review. A review in 2011 by the SAB provided the EPA with four options for combining mortality risk valuation estimates. These options were: (1) Develop independent estimates for relevant cases using only studies that are closely matched on risk and individual characteristics; (2) Develop a baseline distribution of estimates (perhaps for fatal injury) and a set of adjustment factors for risk and individual characteristics as warranted; (3) Develop a meta-regression model to estimate the VSL as a function of risk and individual characteristics; and (4) Develop and estimate a structural preference function (EPA SAB 2011). The White paper notes that these options were evaluated and a composite of two of the four was adopted in developing the analysis presented to the SAB in the White Paper and other documents. More specifically, the White Paper notes that:

> "…in light of the number of studies and estimates that meet the selection criteria recommended by the SAB Environmental Economics Advisory Committee described above, the EPA chose an approach for updating the VSL that blends options 2 and 3. Specifically, we used meta-analysis to estimate the average value (among the U.S. general adult population) of the marginal willingness to pay to reduce the risk of immediate death, hereafter referred to as the VSL. In addition to the meta-analysis, we also estimated a parsimonious meta-regression model that pools all of the observations in the meta-analysis data set and controls for study type (HW or SP), means versus medians, and year of data collection. W leave the task of estimating adjustment factors to account for the influence of risk and individual characteristics on the VSL, possibly through inclusion of additional control variables in the meta-regression model, for future work."

The EPA's composite strategy represents an innovation in benefits transfer practices. By embedding assumptions that adjust estimates of risk tradeoffs for different members of the population based on what is available in the literature and then assigning these values to groups as part of constructing national averages (often for different years than when they were estimated), the EPA has developed a new strategy for introducing heterogeneity into the logic used to construct unit value transfers. This logic also identifies another issue to be considered in adjusting estimates from the literature. To the extent the analysis acknowledges heterogeneity in measures of risk tradeoffs for different groups distinguished by observable characteristics, such as age or gender, then adjustments for increases in income or other sources of risk change that might affect the baseline risk need to be considered prior to constructing a national average measure for VSL. These groups may face different rates of growth in their income or have different income elasticities. They might experience other differences in factors that would affect the risk tradeoff measure relevant for this subgroup, and these would need to be applied before developing a composite nationwide average measure.

The White Paper describes interesting approaches that were not envisioned in the earlier SAB recommendations. It is important to note that one cannot evaluate the final estimates for a national average without evaluating all the assumptions applied in developing those estimates and evaluating whether specific assumptions are influential.

The EPA asked the SAB to review the White Paper and other documents and respond to 17~~19~~ charge questions organized into six topics focusing on~~that focused on~~: (1) whether the methods used to select the data set for the analysis were appropriate and scientifically sound, (2) whether relevant empirical studies were adequately captured in the White Paper, (3) whether the methodology used in the White Paper to analyze the data represents an appropriate and scientifically sound application of meta-analytic methods to derive VSL estimates; (4) whether the EPA's proposed VSL estimates represent reasonable and scientifically sound conclusions; (5) development of a protocol for future updates of the VSL; and (6) whether EPA's approach to estimating the income elasticity of VSL was appropriate and scientifically sound. In response to the EPA's request, the SAB convened its Environmental Economics Advisory Committee to conduct the review. The Committee held a public meeting on March 7-8, 2016 to deliberate on the charge questions and develop a consensus report of its findings and recommendations. This SAB report provides the findings and recommendations of the SAB in response to the EPA charge questions (Appendix A.). The SAB recommendations are highlighted at the end of each section of this report.

Commented [K32]: Before being implemented for policy analysis it should be subjected to rigorous discussion and peer review beyond this committee—where the details of the logic can be fully vetted

Commented [K33]: This statement to me implies that we CANNOT suggest what has been done is standard practice. It is a proposed innovation to standard practice that is interesting but NOT ready to provide the basis for the development of a VSL for use in RIA's. The VSL is the most important benefit number sued by EPA and as such major changes in practice should be completely documented and evaluated before being used for policy. This runs the risk of discrediting the VSL measure.

3. RESPONSES TO EPA'S CHARGE QUESTIONS

3.1. Meta-Analysis Dataset

3.1.1. Evidence of Validity of the Stated Preference Studies

Charge Question 1a. Evidence of validity for stated preference studies: The SAB noted in its earlier advisory report (U.S. EPA Science Advisory Board 2011) that each selected stated preference study "should provide evidence that it yields valid estimates" (page 16). The SAB did not, however, specify how validity should be assessed. In applying this criteria, EPA included studies and estimates that passed a weak scope test or provided other evidence of validity (e.g., a positive coefficient on the risk variable as in the appendix for Viscusi, Huber and Bell 2014) as explained in Appendix B of the White Paper. Please comment on whether the methods EPA used in the White Paper to assess the validity of studies and estimates are appropriate and scientifically sound.

The SAB previously recommended that specific criteria be used in identifying appropriate stated-preference studies to estimate the Value of Statistical Life (EPA SAB 2011). In particular, the SAB recommended that the EPA use only estimates with evidence of validity, such as passing a scope test (i.e., showing that willingness to pay increases with the size of risk reduction within or between samples of respondents in a stated-preference study). The EPA indicated that it applied the SAB's recommended criteria in selecting the studies used in the 2016 White Paper and has asked the SAB to comment on whether the methods used to assess the validity of the studies and estimates are appropriate and scientifically sound.

The SAB finds that ~~that~~ the evidence of validity considered by the EPA in selecting studies for use in the 2016 White Paper is ~~appropriate but~~ incomplete. The following aspects of the methodology for assessing validity should be clarified:

1) Application of the methods to assess study validity. It is not clear how the EPA applied the evidence of validity across all studies included in the analysis and whether the same criteria were applied to all studies (both the included and excluded studies).

2) The list of factors considered as evidence of validity is incomplete, especially with regard to study design and administration features. In order to strengthen the assessment of study validity and better inform a weight of evidence decision to include or exclude a study, the SAB recommends that the EPA expand the consideration of evidence of validity to include answers to the additional key questions discussed below.

3) Threshold for inclusion or exclusion of studies. It is not clear how, or if, evidence of validity was used to exclude or include studies and ~~observations~~ value estimates in the data set used by EPA. Excluded studies and ~~observations~~ value estimates are identified in the White Paper, but all of the exclusions may not be justified. The white paper appendix that discusses assessment of validity (Appendix B) is silent on the investigations of validity for some studies. In addition, the threshold for inclusion of studies and ~~observations~~ value estimates is not clearly stated in the White Paper. This is not a bright line decision, but a consideration of the weight of evidence as

we discuss below. Thus, it is crucially important that EPA fully document the validity evidence considered and how this evidence was used to include each study or value estimate.

4) Further, EPA does not document characteristics of included studies to show that all VSL estimates are estimated using a common metric. Nor are the data manipulations employed to transform value estimates to compute the updated VSL clearly documented. This document needs to be transparent so the analyses are replicable and clear justifications need to be presented for all data manipulations that are supported by economic theory, previous research, or the estimation procedures used.

Addressing many of the concerns raised in this section will require a series of carefully crafted appendix tables where the findings in the appendices are clearly integrated into the main text of the White Paper.

The charge question addressed in this section refers to "stated-preference studies," but the SAB strongly feels that the general considerations discussed in this section apply equally to the revealed-preference studies used in the analysis to update the VSL and the general recommendations should be applied to these studies/value estimates as well.

Study Validity

Validity is not based on a bright-line, valid/invalid criterion. In fact, there are three components of validity and multiple considerations with each component of validity. Thus, the validity of any study or value estimate for inclusion in the analysis to update the VSL must be based on a weight of evidence consideration of features that support a conclusion of validity or invalidity (Bishop and Boyle, 2016).

The three important concepts of stated preference study validity should be considered – content, construct and criterion (Carmines and Zeller 1979). Content validity ~~takes into~~ considers the extent to which a study ~~ation the~~ uses ~~of~~ established procedures to estimate values~~implement a method~~ (e.g., U.S. EPA's *Guidelines for Preparing Economic Analyses*); construct validity involves ~~the~~ testing of responses to valuation questions to investigate if they conform with hypothesized relationships ~~specific procedures~~ (e.g., procedural invariance, convergent validity, tests of scope, etc.); and criterion validity ~~involves~~ investigates if value estimates are statistically the same as an estimate ~~comparison of empirical outcomes against a~~ of the presumed ~~truth~~ true value (e.g., comparisons with cash transactions). Each of these types of validity apply to all types of empirical estimates, including the hedonic, revealed-preference estimates of VSL. The discussion here is customized to the estimation of nonmarket values.

All studies should consider the validity of the resulting value estimates, but no single validity investigation indicates a value estimate is valid or invalid. Further, t~~T~~here is no perfect study and all empirical estimates likely contain some bias; the presence of bias does not, by itself, indicate a value estimate is invalid~~no absolute test of validity~~. ~~Content validity only implies that the likelihood of unbiased estimates is enhanced or, perhaps, that bias is reduced. Satisfying construct validity establishes the credibility of a specific procedure in implementing a stated-preference study. Like content validity, construct validity implies that the likelihood of unbiased estimates is enhanced or that bias is reduced. Criterion validity is the strongest concept of validity as it speaks directly to bias, but the outcome is only as credible as the credibility of the criterion, the measure of the presumed truth. Thus, v~~Validity assessment ~~depends on~~requires consideration of~~ing~~ the weight of evidence regarding content, construct

and criterion validity and is a matter of judgment based on the weight of evidence. This weight of evidence can include investigations conducted as part of the study that generated a value estimate and can rely on evidence published in the peer-reviewed literature. The SAB suggests some criteria for weighing the such evidence of validity below.

The evidence of validity considered in the current draft of the White Paper includes sensitivity to scope and question ordering effects (i.e., the order of presentation of valuation questions in the stated preference survey should not affect responses and corresponding willingness to pay estimates). These two types of validity investigations are good examples of why any decision on validity requires a careful consideration of the weight of evidence. As will be explained below, failure to find a statistically significant scope effect and the presence of a question ordering effect are not, by themselves, evidence of invalidity.

~~However~~Further, it is not clear whether the EPA ~~applied~~ considered these validity tests ~~to~~ for all studies used in the analysis or just to those studies where such evidence was made available by the authors. Therefore, the SAB recommends that in the White Paper the EPA provide a table that lists the evidence of validity that was available (or not) for each of the studies excluded from and included in the agency's analysis. The EPA should document in this table whether ~~the~~ such evidence of validity was used to support exclusion or inclusion of studies and ~~observations~~ value estimates within studies.

Evidence of Study Validity

Scope and question ordering effects ~~fall under the concept~~ are examples of construct validity investigations. Evidence of scope and the lack of a question ordering effect (procedural invariance) are evidence of validity. It is logical to expect willingness to pay to increase for a larger reduction in risk and one would not expect value estimates to change within an arbitrary ~~the~~ sequence of ~~where a~~ stated-preference questions within~~was placed in~~ a survey, but this need not be the case for either of these investigations.

A scope failure or a question sequencing effect, while reason for concern, does not mean a value estimate is invalid. Tests of scope imply structure on respondent preferences that are imposed by the investigators and that may or may not be true (Van Houtven et al. 2011). The basic axioms of choice only imply that marginal willingness to pay for a larger reduction in risk should be nonnegative, not that it should be positive and significantly different from zero or further, proportional to the risk change. Thus, excluding studies with an insignificant scope effect may lead to an overestimation of value. Violation of procedural invariance, due to a question ordering effect, does not establish whether one or both value estimates are biased. In fact, both estimates could be valid within the sequence that the questions were asked and procedural invariance would not be expected to apply (Carson, Flores and Hanemann 1998). Thus, failure of a test of construct validity typically requires additional investigation to understand if the failure is evidence of invalidity or validity ~~that the stated preference estimate is valid and unbiased~~.

~~Further,~~ N~~n~~ot every study ~~needs to conduct nor does~~ conducts or, perhaps, reports ~~a construct~~ validity investigation. This makes it difficult, if not impossible, to consistently evaluate every stated-~~-~~preference study for evidence of validity based on ~~considerations of construct validity~~ the available documentation. ~~On the other hand, content validity implies considerations that should be elements of every stated~~

~~preference question (Boyle 2003). Thus, evidence of validity should consider both content and construct validity. A test of criterion validity is unlikely in most studies and may be impossible in many applied studies. Therefore, criterion validity is not recommended as a required element of evidence of study validity.~~

In order to inform ~~strengthen the assessment of study validity and better inform a~~ weight~~-~~of~~-~~evidence decisions for study/estimate ~~to includ~~inclusion/~~e or~~ exclusion~~de a study~~, the SAB recommends that the EPA expand the consideration of evidence of validity to include answers to the following key questions:

- Was the survey pretested using focus groups, one-on-one interviews, or field pretest?
- Was the survey applied to a random sample of a clearly specified population?
- Did the survey clearly define the baseline risk?
- Did the survey clearly explain the change in risk to be valued?
- Was the valuation scenario consequential (payment mandatory and valuation response have a non-zero probability of influencing provision of the item being valued)?
- ~~Was the sample of respondents investigated for comparability to the population sampled?~~
- Was the stated preference question a binary choice framed as a referendum or product purchase~~ where payment would be required if the risk change was provided~~?
- Were robustness checks conducted of the statistical analyses that led to the value estimate?
- Were construct validity tests conducted?
- Was the sample of respondents investigated for comparability to the population sampled?
- Was the study published in a peer-reviewed journal~~ or evaluated through a transparent and rigorous peer review~~
- ~~process~~?

The first ~~seven~~ six items relate to content validity~~ and should be elements in any credible stated-preference study~~. Construct validity tests (items 7, 8 and 9) provide evidence of validity in one or more dimensions of the study design and implementation. Construct validity could involve any tests of respondents' understanding of the risk scenario and choice they are being asked to value, as well as modeling assumptions imposed by the analyst. Peer review~~, not just in an academic journal (e.g., M.S. and Ph.D. theses have peer review by graduate students' committees),~~ is evidence ~~that~~of the scientific validity of a study~~ has been investigated by one or more peers~~. This broader consideration of validity can inform the weight of evidence supporting the exclusion or inclusion of studies and individual value estimates in the meta-analysis (Bishop and Boyle 2016). Consistent with arguments presented earlier, a study or value estimates need not satisfy every item in the above list to be deemed valid and worthy of inclusion in EPA's analyses. In fact, it may not be possible to determine that a study or estimate is valid, but it may be possible to decide that there is insufficient evidence to support a conclusion of invalidity and the data are therefore worthy of inclusion in the analyses.

The SAB notes that EPA should develop and use a similar set of criteria for evaluation Hedonic estimates of VSL for inclusion/exclusion in the analyses.

Judging the Weight of Evidence

Validity decisions need to proceed with caution. For example, the elements of content validity may or may not be reported in a journal article. As noted above, tests of construct validity are not a prerequisite of any individual study and failure of construct validity does not necessarily imply invalidity. While journal articles typically include a theoretical or methodological twist that will provide evidence of construct validity, what is reported in journal articles may be constrained by space limitations and the specific focus of the article. A broader consideration of peer reviewed and auxiliary supporting documents may include studies with more policy-relevant value estimates and provide more evidence of validity or invalidity. Thus, decisions on validity need to consider the weight of evidence from the elements in the list above that are documented and available.

There is no precedent in the stated preference literature to establish a standard for what is a valid or invalid stated preference study. The closest analog is the National Oceanic and Atmospheric Administration (NOAA) Blue Ribbon Panel report (Arrow et al. 1993), which lists a large number of validity considerations for contingent valuation (CV) surveys, but does not clearly state that all validity considerations must be met for a study to provide useful information. In fact, the NOAA Panel stated:

> "… we try to lay down a fairly complete set of guidelines compliance with which would define an ideal CV survey. A CV survey does not have to meet each of these guidelines fully in order to qualify as a source of reliable information to a damage assessment process. Many departures from the guidelines or even a single serious deviation would, however, suggest unreliability prima facie."

These guidelines were for studies done for estimating nonuse (or passive use) values to support natural resource damage cases filed in court and where the government can recover the cost of conducting the studies from the responsible party, whereas most valuation studies are conducted for academic purposes or to inform policy decisions where the cost of the studies is a major consideration with hard budget constraints that limit the possible design features. The lack of clear guidance on assessing validity suggests that the EPA needs to proceed cautiously when making validity assessments. Therefore, the SAB recommends careful documentation of studies that meet or do not meet validity criteria as evidenced in the answers to the key questions listed above. ~~Studies that are on the margin for inclusion/exclusion in t~~The meta-analysis should be carefully scrutinized (including sensitivity analyses) to assess whether ~~the potential threats to validity are likely to bias the value estimates~~the inclusion/exclusion of studies and value estimates influences empirical outcomes.

Other Validity Assessments

The discussion above has focused on whether a specific stated preference study or ~~observations withi~~value estimate from~~n~~ a study are valid. There are also broader validity assessments that can be conducted to determine whether a body of literature is valid and whether a method is valid. With regard to method validity, there can be evidence in the literature that establishes whether stated-~~-~~preference design and implementation procedures lead to valid value estimates. These general insights from the literature can provide evidence~~us, while a specific study may not provide evidence of investigation of a specific design or implementation feature, it is possible to refer to the broad valuation literature to assess of~~ validity or invalidity. Following the discussion of a consequential valuation scenario, a study that contained the key elements of consequentiality might be deemed valid based on the understood wisdom in the peer-reviewed literature even if the individual study did not conduct a statistical investigation of

consequentiality. ~~Validity of a body of empirical research, such as VSL estimates, can be investigated through meta-regressions to consider points of validity and invalidity (Mrozek and Taylor 2002; Lindhjem et al. 2011)~~. The SAB recommends that weight of evidence assessments of study validity should be informed by consideration of the broad stated~~-~~-preference literature and pre-existing meta-analyses of VSL (e.g., Mrozek and Taylor 2002; Lindhjem et al. 2011) .

Updating the VSL Estimate

It is important that the knowledge and assessment of study validity evolve through time as research progresses. Future updates of the VSL should consider advancements in the literature pertaining to study design, conduct, and testing relating to validity. ~~An~~ For example, ~~of this is the current evolution in the literature extending incentive compatibility through~~a consequential survey design (Carson, Groves and List 2015) was not a central point of discussion in the last update of the VSL, but is a study-design component that should be considered in any contemporary evaluation of a stated-preference study.

Such updating does not necessarily exclude older studies. For example, while consequentiality has only entered the stated preference literature in recent years, many, if not most, earlier studies following good practices were consequential. One way to consider consequentiality is to contact investigators and request their survey instruments to evaluate if a binding payment was used and if subjects were informed that their responses would influence provision of the item being valued. Many earlier studies may not have included questions to inquire if subjects considered the valuation exercise consequential so this would be much harder to assess from an ex ante perspective.

Therefore, the SAB recommends that all future updates of the VSL simultaneously consider whether the conditions for investigating study validity should be updated. The recommendations here are for processes to follow and not hard and fast rules that are invariant over time.

Identification of all Criteria for Study Inclusion/Exclusion

The SAB finds that the White Paper could be improved by ~~identifying~~transparent documentation of:

(1) ~~all~~ criteria for including studies in the meta-analysis,
(2) characteristics of included studies, and
(~~2~~3) ~~all manipulations of value estimates that were performed to convert the estimates to a homogenous metric that would support the meta-analysis (e.g., manipulation of the value estimates in a study that has a well-defined baseline risk and risk change but is not consistent with the other studies that are included in the meta-analysis)~~ all assumptions and data manipulations affecting the credibility of the final VSL estimate, e.g., adjustment by income elasticity, mixing of Marshallian, willingness to accept (WTA) a risk increase, and Hicksian, willingness to pay (WTP) for a risk decrease measures, discounting assumptions to get an annual value, etc.

The first item has already been extensively discussed. The second item is important because study design features that are not addressed in the validity considerations can affect the magnitude of VSL estimates. For~~An~~ example, ~~that does not affect validity, but affects the measure of values and the computation of VSLs, is~~ the time frame of payment (e.g., one time, annual payment for a fixed period, continuing payments into perpetuity) can substantially affect VSL estimates in a stated-preference study.

Such differences across studies can lead to manipulations of actual VSL estimates to convert the estimates to a common metric for the meta-analysis, which leads to the third item above. Such a data manipulation and other data manipulations, such as the examples listed with the third item, ~~Such criteria and manipulations of value estimates~~ should be identified, clearly justified and presented in a transparent manner to allow the data manipulations and analyses to be replicated. ~~even though they may not be fundamental to investigation of the validity of VSL estimates~~. ~~An enhanced listing of study features considered for inclusion/exclusion of studies and observations, and a listing of the manipulations of the reported value estimates, would avoid confusion between validity of the underlying values and validity of the~~Addressing these three points enhances the validity and credibility of the meta-analysis itself and the resulting VSL estimate to support policy analyses.

Therefore, the SAB recommends that all criteria for inclusion/exclusion of studies/~~observations~~ value estimates be documented systematically across studies in a table similar to the one previously recommended for documenting the investigation of study validity. This should be followed by ~~a~~ separate tables that document~~s~~ included study characteristics and data manipulations ~~of value estimates to convert value estimates to a homogenous metric to support the meta-analysis~~. This will help to ensure that the VSL data and analyses are transparent and replicable. The study characteristics should contain enough information to allow readers to assess how the reliance on published studies, particularly other meta-analyses (including studies that drew from international data) might lead to results that differ due to publication bias, lags in publication, data sources included, methodology relied upon, or other concerns.

Key Recommendations

- In the White Paper the EPA should provide a table that lists the evidence of validity that was or was not available for each of the studies considered for inclusion in the agency's analysis. The EPA should also document in this table ~~whether~~ how the evidence of validity was used to support exclusion or inclusion of studies and ~~observations~~ value estimates within studies. The White paper should also clearly indicate the types of studies other than hedonic wage or stated preferences that were available for use but eliminated by screening criteria. EPA should provide a rationale for excluding these types of studies.

- Consideration of evidence of study validity should be expanded to include answers to the following questions:

  - Was the survey pretested using focus groups, one-on-one interviews, or field pretest?
  - Was the survey applied to a random sample of a clearly specified population?
  - Did the survey clearly define the baseline risk?
  - Did the survey clearly explain the change in risk to be valued?
  - Was the valuation scenario consequential (payment mandatory and valuation response have a non-zero probability of influencing provision of the item being valued)?
  - Was the stated preference question a binary choice framed as a referendum or product purchase?
  - Were robustness checks conducted of the statistical analyses that led to the value estimate?
  - Were construct validity tests conducted?
  - Was the sample of respondents investigated for comparability to the population sampled?

- Was the study published in a peer-reviewed journal~~ or subject to peer review~~?
- ~~Was the survey pretested using focus groups, one-on-one interviews, or field pretest?~~
- ~~Was the survey applied to a random sample of a clearly specified population?~~
- ~~Did the survey clearly define the baseline risk?~~
- ~~Did the survey clearly explain the change in risk to be valued?~~
- ~~Was the sample of respondents investigated for comparability to the population sampled?~~
- ~~Was the stated-preference question a binary choice framed as a referendum or product purchase where payment would be required if the risk change was provided?~~
- ~~Were robustness checks conducted of the statistical analyses that led to the value estimate?~~
- ~~Were construct validity tests conducted?~~
- ~~Was the study published in a peer-reviewed journal or evaluated through a transparent and rigorous peer review process?~~

- ~~The EPA should clearly document all of the reasons why the included and excluded studies have or have not met validity criteria. Studies that are on the margin for inclusion/exclusion in the meta-analysis should be carefully scrutinized (including sensitivity analyses) to assess whether the potential threats to validity are likely to bias the value estimates.~~

- Evidence in the literature can be used to establish whether stated-~~-~~preference study design and implementation procedures lead to valid value estimates. That is, validity need not be solely based on statistical investigations conducted within studies, but also on considerations of the consistency of studies with design, implementation and analysis features that are consistent with established practices in the peer reviewed literature. For, example, insights on the v~~V~~alidity of a body of empirical research, such as VSL estimates, can be investigated through meta-regressions. Therefore, weight of evidence assessments of study validity should be informed by consideration of the broad stated-~~-~~preference literature and pre-existing meta-analyses of VSL.

- ~~All future updates of the VSL should simultaneously consider whether the conditions for investigating study validity should be updated.~~

- All criteria for inclusion/exclusion of studies/~~observations~~ value estimates should be documented systematically for each ~~across~~ study considered, and characteristics of included studies should also be documented along with all data manipulations used to adjust the data for the update analysis of the VSL~~ies in a table in the White Paper~~. ~~This should be followed by a separate table that documents manipulation of value estimates to convert the estimates to a homogenous metric to support the meta-analysis.~~

- The peer-reviewed literature is not static. All future updates of the VSL should simultaneously consider whether the conditions for investigating study validity and meta-analysis procedures should be updated, not just the VSL itself. For example, at the time of the last VSL review, consequentiality was not a predominant feature for evaluating study validity and it was not standard practice to investigate the sensitivity of meta-analysis estimates to study and value estimate inclusion/exclusion in the estimation. These are important, new features of the VSL updating today and there are likely to be other new innovations in the literature at the time of the next updating of the VSL.

3.1.2. Construct of the Risk Variable in Hedonic Wage Studies

Charge Question 1b. Construct of the risk variable in hedonic wage studies: The SAB noted in its earlier advisory that the EPA should "Eliminate any study that relies on risk measures constructed at the industry level only (not by occupation within an industry)" (U.S. EPA Science Advisory Board 2011, page 18). It is not clear whether the SAB's parenthetical addition was meant as an example or as a directive. Only four studies constructed the risk variable by occupation and industry and met other selection criteria. In applying this criteria EPA included studies and estimates where the risk measure is differentiated by industry and at least one other characteristic (e.g., occupation, gender, age). Please comment on whether the hedonic wage studies included in the White Paper constructed the risk variable in a manner appropriate for use in the meta-analysis.

The ~~VSL~~SAB supports excluding from the analysis ~~in EPA's White Paper was conducted using hedonic wage studies where the risk measure is~~ those studies that employ fatality risk measures based on industry category alone but finds the current inclusion criterion that restricts the analysis to studies based on risk measures differentiated by industry and "at least one other characteristic ~~(e.g., occupation, gender, age). The SAB finds that the use of "industry and one other characteristic" risk measures is not appropriate for characteristics such as gender or age. This is because the risks are likely to be too varied within~~" inappropriate. Further differentiating an industry ~~and because~~ level risk measure by some additional characteristics, for example age and gender, is unlikely to sufficiently resolve the measurement error problem and may lead to wage-risk tradeoff estimates unequally influenced by wage discrimination.

The SAB ~~might affect the wage-risk differential across some groups. Therefore~~makes the following two short run recommendations to EPA. First, the SAB recommends that the EPA alter its inclusion criterion to restrict its analysis to ~~determine the final VSL summary measure only include~~ hedonic wage studies ~~where the~~that employ fatality risk ~~variable includes variation~~measures differentiated by occupation. This change would lead to the inclusion of studies that use risk measures based on occupation ~~– either with respect to occupation and industry or with respect to occupation only. One example of an "occupation-only" hedonic wage study that should be included in the analysis is Delaire~~alone (e.g., Deleire, Khan, and Timmins 2013) as well as those studies based on risk measures differentiated by occupation and another characteristic, such as industry (e.g., Viscusi 2004). Second, the SAB ~~also~~ recommends that the EPA include in the White Paper a summary of recent meta-analyses of hedonic wage studies. The summary should provide information about how the results of those studies vary according to study design and data sources (e.g., alternative risk measures, studies without a morbidity risk measure, sub-national geography within the U.S., and possibly studies from other countries). This summary will enable the White Paper to convey the likely sensitivity of the final VSL summary measure to variations in the set of studies included in the calculations without having to ~~repeat the work already done~~replicate the research efforts completed in those meta-analyses. Recent VSL meta-analyses include: Miller (2000); Bowland and Beghin (2001); Mrozek and Taylor (2002); De Blaeij et al. (2003); Viscusi and Aldy (2003); Kochi et al (2006); Robinson (2008); Bellavance et al. (2009); and OECD (2012).

Enhancing the Quality of VSL Estimates Generated by Hedonic Wage Studies

The SAB finds that research is needed to enhance the quality of VSL estimates generated by hedonic wage studies. ~~It is recommended~~The SAB makes two long run recommendations toward this end. First, because publication in primary economics journals favors innovation in methods rather than replication, the SAB recommends that the EPA compile, make publicly available (e.g., on an internet webpage), and regularly update data that would encourage future revealed preference VSL research. To this end, the SAB recommends that the EPA pursue ~~research in the areas discussed below. The agency should examine existing research and either conduct~~a Memorandum of Understanding with the U.S. Bureau of Labor Statistics (BLS) to access and, if possible, gain permission to make publicly available a data set that merges appropriate data from the Current Population Survey (CPS) and the Census of Fatal Occupational Injuries (CFOI). The CFOI data required to create fatality risk measures differentiated by occupation and industry are currently available only to those researchers with BLS data agreements. By hosting the merged data on an EPA platform, the agency could resolve some of the challenges in developing these studies (e.g., accessing the occupation and industry differentiated risk measures), signal the importance of new research ~~or encourage others to undertake research to~~, offer a resource for innovative benefit transfer using these data (independent of new estimates of hedonic wage models), and enhance the quality of future updates of the VSL.

Commented [K34]: It was also provide a valuable data base for potential use by EPA or its contractors in establishing estimates of baseline risks for policy evaluation associated with RIAs.

~~The SAB is~~Second, the SAB recommends that the EPA use the above-mentioned data to apply a consistent hedonic wage model to all of the available years to generate comparable annual measures of VSL. Comparisons among these annual VSL estimates yield an estimate of the income elasticity of the VSL (IEVSL). The SAB recommends that this analysis be regularly updated as new data become available. This analysis would be relatively inexpensive to conduct, could be done by EPA staff or by other researchers, and would assist EPA in systematically updating its estimates of the VSL and IEVSL over time.

In addition to these four recommendations, the SAB identifies several areas of research that would improve the quality of future updates of the VSL. The first broad category includes research that examines the potential biases associated with hedonic wage studies. The SAB notes that the VSL generated by hedonic wage studies may be systematically biased relative to the VSL for the whole population and may differ from the VSL obtained from stated preference studies. The following four potential ~~biases~~issues are noted. ~~These are evenly split between upward and downward biases.~~

1. Limited worker awareness of risks or limited worker mobility across jobs could lead the hedonic VSL to understate workers' true preferences. Worker misperception of risk could also lead to underestimation of the VSL.

2. Sorting of risk-averse workers into safer jobs could lead the hedonic VSL to understate the average preferences of the whole population.

3. The hedonic sample includes only ~~workers. They~~employed individuals, who may have ~~higher income~~ systematically different risk preferences than the rest of the population ~~and this~~ due to, for example, higher incomes. These selection effects could lead the hedonic ~~VSL~~wage model to overstate or understate the average preferences toward risk reductions for the whole population.

4. ~~The hedonic VSL is related to a Marshallian willingness to accept which could overstte the Hicksian willingness to pay measure.~~ The VSL estimate from hedonic wage studies relates to an estimate of the marginal rate of substitution that, under ideal conditions, holds non-wage income constant. As such it can be interpreted as a measure of Marshallian willingness to accept (WTA) a marginal increase in fatality risk. The relationship between this WTA measure and the Hicksian willingness to pay (WTP) for a marginal reduction in fatality risk is not well understood. In general, we expect WTA measures to be larger than WTP measures. However, in this setting, it is difficult to state *a priori* the direction of the relationship. The SAB notes that ~~a simple algebraic formula (and some assumptions) could identify how large an adjustment would be needed to convert the Marshallian measure to the Hicksian one. This~~ the approach ~~is~~ described in Smith~~,~~ et~~.~~ al. (2006~~). An assessment of~~) could be used to assess the magnitude of ~~this adjustment could be conducted immediately~~the difference between these two measures in order to determine the importance of this issue.

~~The SAB also notes that existing hedonic wage studies are often based on different models or data from different sources and therefore it may be difficult to find future published hedonic wage studies that carefully apply existing hedonic wage models to new data. The EPA should consider applying a consistent hedonic wage model to the available years of data, combining an industry and occupation risk measure from the U.S. Bureau of Labor Statistics Census of Fatal Occupational Injuries (CFOI) with the U.S. Bureau of Labor Statistics March Current Population Survey (CPS) wage information, and generating measures of VSL on a consistent basis. Estimates for future years should be added as the data become available. This research would be relatively inexpensive to conduct, and could be done by EPA staff or by other researchers. This research would assist EPA in systematically updating VSL numbers over time.~~

~~In addition, the SAB notes that differences in estimated VSL across analyses using different risk measures are not well understood. This limits the set of hedonic wage studies that can be included in the VSL calculation.~~A second category includes research that examines the underlying factors that help to explain differences in VSL estimates across hedonic wage studies employing different risk measures. Viscusi (2004) examines both an industry-only risk measure and an industry and occupation risk measure using the same sample and the same model. The results run counter to the classical measurement error model, which predicts lower impacts for the industry-only risk measure because everyone in the same industry is mistakenly assigned the same risk. Instead, the industry-only VSL results are twice as large as the industry and occupation results. Understanding why these differences occur could provide guidance on ~~which combinations of industry and other characteristics would be~~ the appropriate ~~for calculating the~~ risk ~~measure~~measures to consider including in ~~hedonic wage studies.~~future updates of the VSL.

Key Recommendations

- ~~The~~In the short-run, the SAB recommends that the EPA alter its inclusion criterion to restrict its analysis to ~~determine the final VSL summary measure should only include those~~ hedonic wage studies ~~where the~~ that employ fatality risk ~~variable includes variation~~measures differentiated by occupation~~— either with respect to~~. This change would result in the inclusion of studies that use risk measures based on occupation alone (e.g., Deleire, Khan, and ~~industry or with respect to~~Timmins

2013) as well as those studies based on risk measures differentiated by occupation ~~only.~~and another characteristic, such as industry (e.g., Viscusi 2004).

- In the short-run, the SAB recommends that the EPA include in the White Paper a summary of recent meta-analyses of hedonic wage studies.

- ~~EPA should include in the White Paper a summary of recent meta-analyses of hedonic wage studies. The summary should provide information about how the results of those studies vary according to study design and data sources (e.g., alternative risk measures, studies without a morbidity risk measure, sub-national geography within the US, and possibly studies from other countries).~~

- In the long-run, the SAB recommends that the EPA pursue a Memorandum of Understanding with the U.S. Bureau of Labor Statistics to access and, if possible, gain permission to make publicly available appropriate data that consistently merge the records from the Current Population Survey (CPS) with the Census of Fatal Occupational Injuries (CFOI).

- In the long-run, the SAB recommends that the EPA apply a consistent hedonic wage model using data from the CPS and CFOI to generate comparable annual measures of VSL. Comparisons among these annual VSL estimates could yield an estimate of the income elasticity of the VSL (IEVSL). The SAB recommends that this analysis be regularly updated as new data becomes available.

- ~~An~~ The SAB recommends that the EPA pursue, and encourage others to pursue, research to examine the various biases associated with hedonic wage studies including an assessment ~~should be conducted to determine the magnitude~~ of the adjustment needed to convert the ~~hedonic (willingness~~willingness to accept~~) VSL~~ measure obtained from hedonic wage studies to a Hicksian willingness to pay measure.

- ~~A consistent hedonic wage model should be applied to the available years of data, combining an industry and occupation risk measure from the U.S. Bureau of Labor Statistics Census of Fatal Occupational Injuries (CFOI) with the U.S. Bureau of Labor Statistics March Current Population Survey (CPS) wage information, and generating measures of VSL on a consistent basis.~~

- ~~Research should be undertaken to develop a better understanding of differences in estimated VSL across analyses using different risk measures.~~

3.1.3. Estimates of Value of Immediate Risk Reduction

Charge Question 1c. Estimates for immediate risk reductions: To estimate the average value of the marginal willingness to pay for reduced risk of immediate death, the EPA selected estimates from the stated preference literature that are most closely comparable to the accidental deaths from the hedonic wage literature. The EPA made several judgment calls in determining the appropriate estimates to use from the stated preference literature. Specifically, Viscusi, Huber and Bell (2014) estimate reductions in risk of bladder cancer that will occur in 10 years. The authors discount the estimates to derive a comparable estimate for an immediate risk reduction. Alberini, et al. (2004) estimate a willingness to pay for an annual reduction in risk over 10

years. We include estimates from both of these studies in the meta-analysis. Please comment on whether appropriate estimates from the stated preference literature were used in the White Paper to estimate the marginal willingness to pay for reduced risk of immediate death.

The SAB was asked to comment on whether the agency selected appropriate estimates from the stated preference literature for its analysis of willingness to pay for reduced risk of immediate death. As discussed in Section 3.1.4 of this report, the SAB has provided citations for several additional VSL studies that could be included in the White Paper. In addition, the SAB finds that the supplementary analysis in one of the studies selected by the EPA for use, Viscusi, Huber, and Bell (2014), does not provide clear evidence of study validity (i.e., sensitivity of scope).

Use of a Benefits-Transfer Approach

Moreover, the SAB finds that, as in other areas of environmental valuation, the limited available VSL literature points to use of a benefits transfer approach. EPA should consider using best practice benefits-transfer methods that employ principled adjustments in existing estimates to fit the particular policy problem of interest. In contrast, meta-analysis relies heavily on a statistical weighting of evidence to produce a single value.

The SAB recommends broadening the scope of studies the EPA uses to derive values for reducing both mortality and morbidity risks. There are a significant number of published studies that estimate willingness to pay for improved health and reduced health risks (see studies listed in Appendix B of this report). There also is a burgeoning literature on benefit-risk and risk-risk tradeoff preferences in health and health care that could provide a basis for enriching the evidence base on risk preferences and providing support for benefits-transfer applications (see studies listed in Appendix C of this report). Unlike the expected small increments in the VSL literature over the foreseeable future, there is a strong demand and growing funding for ~~stated~~ stated-preference benefit-risk studies in health and health care as the result of recent U.S. Food and Drug Administration (FDA) regulatory guidance on conducting such studies (U.S. Department of Health and Human Services 2015).

Other Concerns about the Estimation of Willingness to Pay for Reduced Risk of Immediate Death

The SAB also notes the following additional concerns about EPA's general approach to estimation of willingness to pay for reduced risk of immediate death.

1. The risk of immediate death is not a policy-relevant outcome. Virtually all deaths of policy interest occur with latency and are preceded by a period of morbidity and disability, including potential pain and discomfort associated with treatment as well as the ultimately fatal condition itself. It is desirable to distinguish values based on short-term versus long-term effects.

2. ~~Discounting does not correctly account for the effect of time on VSL. Dying immediately means fewer years of life, not just a delay in a financial payment. A more correct construct would be the value of statistical life-years lost rather than the present value of a future statistical death.~~ Simple discounting thus does not account for confounded morbidity values in converting future deaths to equivalent immediate-death values. Cameron and DeShazo

(2013) estimate the stated-preference willingness to pay to reduce risks of different illness profiles. With a discount rate of 3%, willingness to reduce the risk of sudden death corresponds to a willingness to pay of $8.33 million per microrisk (in 2003 U.S. dollars). Illness preceding death adds a morbidity premium, as willingness to pay to prevent one year of sickness before death is valued at $9.22 per microrisk. Gentry and Viscusi (2016) derive a morbidity-component value of up to 25% of total VSL from the Census of Fatal Occupational Injuries data for non-instantaneous deaths.

3. Survey respondents may not be able to precisely evaluate long-latency risks, particularly when there is considerable uncertainty regarding timing of conditions, so value estimates of future risks may be imprecise.

4. EPA used estimates of willingness to pay for reduced risk of immediate death (Viscusi, Huber, and Bell 2014). In this study the authors estimate reductions in risk of bladder cancer that will occur in 10 years. A discount rate of three percent was applied to derive a comparable estimate for an immediate risk reduction. The SAB finds that the selection of a three percent discount rate is arbitrary.
~~and recommends that the EPA use a more correct construct such as the value of statistical life-years lost rather than the present value of a future statistical death.~~

Key Recommendations

- The SAB has provided citations for several additional studies that could be included in the White Paper. In addition, the SAB finds that one of the studies selected for use, Viscusi, Huber, and Bell (2014), does not provide clear evidence of study validity (i.e., sensitivity to scope).

- The SAB recommends that the EPA broaden the scope of studies used to derive values for reducing both mortality and morbidity risks. There are a significant number of published studies that estimate willingness to pay for improved health and reduced health risks, ~~and~~ a burgeoning literature on benefit-risk and risk-risk tradeoff preferences in health and health care, and transportation literature on reduced risk for highway fatalities that could provide a basis for enriching the evidence base on risk preferences and providing support for benefits-transfer applications.

- ~~Discounting does not correctly account for the effect of time on VSL. The SAB recommends that EPA use a more correct construct such as the value of statistical life-years lost rather than the present value of a future statistical death.~~

- The SAB recommends that EPA account for morbidity values in converting future mortality risks to equivalent instantaneous risks.

3.1.4. Empirical Studies

Charge Question 2. Please comment on whether relevant empirical studies in the stated preference and hedonic wage literatures are adequately captured in the White Paper. If additional studies should be included in the white Paper please provide citations.

The SAB was asked to comment on whether relevant empirical studies in the stated preference and hedonic wage literatures are adequately captured in the White Paper. The SAB finds that there has been a lack of significant growth in the number of studies used by EPA to estimate the VSL ~~literature~~ since the last consideration of this topic by the SAB in 2011. The SAB recommends that EPA search more broadly for additional studies not restricted to Hedonic or stated preference methods, such as evaluating whether studies using experimental or quasi-experimental methods may exist that offer insight to VSL. The SAB suggests that the EPA consider the following additional VSL studies: Ashenfelter and Greenstone (2004); Davis (2004); Deleire, Khan, and Timmins (2013); Viscusi and Gentry (2015); and Gentry and Viscusi (2016). However, no firm recommendations are provided on how these studies might be incorporated in the White Paper, if at all. The SAB also suggests that the EPA consider hedonic studies other than those related to hedonic wage rates. ~~Therefore, t~~The EPA may need to commission more studies or create other incentives for new studies in order to improve the prospect for a deeper literature to support future reviews of VSL. ~~The SAB suggests that the EPA consider the following additional VSL studies: Ashenfelter and Greenstone (2004); Davis (2004); Deleire, Khan, and Timmins (2013); and Viscusi and Gentry (2015). However, no firm recommendations are provided on how these studies might be incorporated in the White Paper, if at all. The SAB also suggests that the EPA consider hedonic studies other than those related to hedonic wage rates.~~

In considering whether relevant studies are adequately captured in the White Paper, it is important to recognize a number of limitations related to the scope of hedonic wage studies, particularly in relation to forms of sampling bias and the ability of these studies to provide a nationally representative estimate in the absence of assumptions needed to extrapolate from subpopulations included in published studies to a broad national population. In particular, as previously indicated, hedonic wage studies exclude non-workers, so the EPA should address the implications of using studies that fail to address individuals' choices of whether or not to work, rather than a near-exclusive focus on valuation derived from choice among different jobs with different risk levels. The SAB suggests that the EPA consider using hedonic wage studies that apply data other than the CFOI data, while acknowledging concerns that studies based on survey data may be subject to non-response biases. The SAB views the quality of the CFOI data as representing the minimum quality of data that should be considered, while recommending that EPA identify other data sources of similar or higher quality that may provide a valid foundation for estimation of VSL~~Studies that use data of lower quality than the CFOI data should be excluded.~~

The SAB also provides specific recommendations concerning clarification of the study selection process and potential limitations of studies used in the White Paper. The White Paper should contain more detail or information, likely in appendices, to allow readers to assess how the reliance on published studies, particularly other meta-analyses (including studies that drew from international data), might lead to results that differ due to publication bias, lags in publication, data sources included, methodology relied upon, or other concerns. Additional information is needed in the White Paper to more clearly indicate the types of studies, other than hedonic wage or stated preferences that were available for use but eliminated by screening criteria. The SAB notes that existing meta-analysis studies might provide insight into the foundation for maintaining or altering study screening criteria.

Key Recommendations

- The EPA should consider commissioning more studies or creating other incentives for new studies in order to improve the prospect for a deeper literature to support future reviews of VSL.

- EPA should consider including the following additional VSL studies in the White Paper: Ashenfelter and Greenstone (2004); Davis (2004); Deleire, Khan, and Timmins (2013); and Viscusi and Gentry (2015)~~.~~; and Gentry and Viscusi 2016). While these studies differ in methodology, data, or approach from studies already included, the SAB believes they offer a potentially valid insight to the estimation of VSL. If excluded, EPA should provide a justification for exclusion of any studies so excluded.

- In the White Paper the EPA should address limitations of hedonic wage studies, particularly in relation to forms of sampling bias and the ability of these studies to provide a nationally representative estimate in the absence of assumptions needed to extrapolate from subpopulations included in published studies to a broad national population.

- ~~The White Paper should contain more detail or information to allow readers to assess how the reliance on published studies, particularly other meta-analyses (including studies that drew from international data), might lead to results that differ due to publication bias, lags in publication, or other concerns *<<Chair's note: a sentence could be added to suggest how EPA should do this.>>*~~

- ~~The White Paper should more clearly indicate the types of studies, other than hedonic wage or stated preferences that were available for use but eliminated by screening criteria.~~

3.1.5. Population Weighting in EPA's Analysis

Charge Question 3. Some estimates in the meta-analysis dataset in the White Paper are constructed by weighting subpopulation-specific estimates within a study in order to approximate an estimate for the general population. The specific weights used are described in Appendix B of the White Paper. Please comment on whether the population-weighting approach used in the White Paper is appropriate and scientifically sound.

The SAB was asked to comment on whether the population weighting approach used in the White Paper to approximate a VSL estimate for the general population is appropriate and scientifically sound. EPA's study screening criteria include selection of studies that provide a national level representation of general U.S. adult population. However, it is clear from the detailed notes in the White Paper (Appendix B) on the selection and weighting of estimates from each study included in the meta-dataset that the agency necessarily incorporated studies that met this criterion only "sufficiently" (i.e., some estimates were reported by subgroup and these were weighted to derive a population estimate). As a result, several of the studies ~~appear to omit~~ draw on data ~~directly applicable to~~that do not include some portion of the U.S. adult population. For example, hedonic wage studies appear to omit non-workers, which tends also to omit older and younger individuals~~ above a standard that each study set~~. This raises a question. Does the assumption that different demographic categories of individuals have different VSLs reflect fundamentally different risk preferences, or is there a cohort effect whereby as young people age, their preferences become more like older individuals? EPA also included o~~O~~ther hedonic wage studies which focused on ~~used selected~~ specific age groups within the workforce, rather than attempting to use data across the full spectrum of workers, for the purposes of the authors of published studies.

General Comments on the EPA's Weighting Approach

The SAB previously recommended that the EPA select studies that are representative of populations affected by EPA regulations (EPA SAB 2011). Given the limited VSL literature, the SAB recognizes the need to develop an ~~weighting~~ approach in order to use subpopulation estimates of VSL in the EPA's analysis. The SAB recommends that EPA provide detail sufficient to: (1) allow a third party to replicate their approach; and (2) distinguish between the use of population weights to derive a representative estimate of the VSL observation drawn from the sample used in a particular study and the strategy (not necessarily using population weights) used to transfer benefit estimates from a source-study as input to an estimate of VSL for some population (or timeframe) not directly addressed by the source study. That is,~~However~~, the White Paper should provide additional explanation of how population weighting was actually done and the strategy used by EPA to bring together observations (or estimates) drawn from several studies to create ~~how the studies were brought together for~~ the aggregate estimate. Adequate information should be provided to enable a reader to replicate the results. In particular, the following issues should be addressed.

1. EPA's approach to weighting, in many cases, focused on deriving an estimated mean. The White Paper should provide a more detailed explanation of how population-weighting -procedures would affect estimates of standard errors, distinguishing the ~~how~~ procedures to derive a nationally representative estimate from one study and procedures to transfer information from a sub-population considered by a particular study for application in another context may have effected estimates and standard errors.

2. Subpopulation weighting ~~Weighting by population shares is common but~~ may not ~~cover~~ acc~~e~~ount for all of the potential sources of selection bias that could result in exclusion of some members of the intended population, either by choice of authors of original studies or by response bias or other known factors contributing to selection bias ~~, particularly for survey-based studies~~. The White Paper should more explicitly address the implications of selection bias that may be present in studies used or excluded. -Furthermore, EPA should distinguish the use of weights to adjust for sampling or response bias on observable characteristics from whether or how EPA has been able to account for selection bias due to unobservable characteristics, such as individuals' risk attitudes.

3. Weighting approaches should ~~to~~ give much greater consideration to details of the specific studies being weighted. Appendix B mixes discussion of two kinds of procedures, population weighting and benefit transfer. In some cases, e.g., contingent valuation studies, the weighting procedure used by the EPA is comparable to using approximations for sampling weights. However, the procedure used for the hedonic wage studies is a benefit transfer. ~~Population weighting and benefit-transfer weighting may involve different principles and relevance.~~ It is important that population weighting be accomplished using standard procedures and that benefit transfer assumptions and procedures for implementation be described and distinguished.

4. As previously mentioned, several of the studies do not provide representation across all possible groups (age, income, employment, ethnicity, agricultural workers, etc.) that necessarily compose

a truly representative sample. The White Paper should discuss the implications of the resulting limitations of source-studies~~his~~ and clearly describe any procedures or calculations that EPA implemented to mitigate these limitations or implications.~~should be discussed in the White Paper.~~

5. Weights should be tied to the time period of the original study, at least for the development of a representative estimate supported by that study, while aggregating available estimates across studies to obtain an overall estimate for 2013. This raises questions of whether weights should correspond to the sample the study is intended to represent or to the full U.S. population.

6. Income adjustments of the VSL estimates derived from hedonic wage and contingent valuation studies must be consistent with the income concept relevant to each model. For hedonic wage models income is endogenous. However, for ~~contingent valuation this~~stated preference studies, this is not the case and expected utility is being held constant. The analysis of proper treatment of income should reconcile the~~se~~ modeling assumptions (including the role of income in Hicksian or Marshallian-based analyses) used by source studies with the use of any approach to deriving an estimate of VSL based on that study, ~~-~~before applying ~~some~~any adjustment.

Commented [K35]: These detailed comments are not consistent with the general conclusion in the summary; this would imply the weighting procedure is NOT necessarily *appropriate and scientifically sound.* They are not doing population weighting as conventionally defined for surveys.

~~Weighting to adjust for income differences in the populations (or time periods) in individual studies should be done after determining the estimates to be drawn from a particular study time period. Income adjustments should be addressed in the process of aggregating across studies to derive an estimate for a representative population. In this regard, there should be an explanation in the White Paper of how Hicksian and Marshallian measures of VSL should be aggregated with a consistent measure of income to account for income effects. *<<Chair's note: it would be helpful to provide a method and citation.>>*~~

Specific Comments on the Weighting of Subpopulation Estimates

The SAB provides specific comments on the manner in which the weighting of particular estimates was conducted in the White Paper analysis to accomplish both the benefit-transfer purpose and the aggregation purpose in supporting the VSL estimate for a representative population.

There are two sets of weights that affect the estimates in the White Paper. The set of weights described in the body of the White Paper concerns the weights applied to the various summary statistics describing VSL estimates from each study. The second set of weights is discussed primarily in Appendix B of the White Paper. It appears that the process discussed in Appendix B was not used in all studies. The first mention of these weights is on page 50 of the White Paper and relates to the Cameron, DeShazo, and Johnson (2010) study where the 28 estimates reported in Tables 4 and 5 of that study were summarized. These estimates are distinguished based on number of children, respondents' gender, age, and marital status. The discussion in the White Paper suggests that the 2010 U.S. Population Census was used to develop a weighted average of the 28 estimates. However, the SAB notes that the Cameron–DeShazo survey was conducted in December 2002 and questions why the EPA did not use the 2000 Census to develop the weighted average. The SAB also notes that the discussion in the background material for the Cameron–DeShazo research indicates that the Knowledge Networks Panel used for the research was representative of the 2000 census. Therefore, Knowledge Network weights could also have been used. When population weighting is necessary to develop an observation of a representative VSL from a source study, SAB recommends that EPA use population weights that are drawn from, for example, the

U.S. Population Census available for a time that is, to the extent practicable, contemporaneous to the data used in a source study.

In addition, the SAB finds that clarification of the weighting process is needed with regard to the following specific issues.

1. Some stated preference surveys used in the White Paper are based on samples but do not report averages for subpopulations. It is not clear whether this is the reason why no weights were applied in these cases.

2. Estimates from the Cameron, DeShazo, and Stiffler (2013) study used in the White Paper were also based on the 2002 samples. Again, 2010 weights were used but the demographic allocation was different. It is not clear whether the weights reconciled since it appears the second set would be an aggregate of the first.

3. The Cameron and DeShazo (2013) study is again based on the same 2002 sample. The weighting approach described on page 55 of the White Paper should be clarified. It notes that "The first four estimates were weighted with each of the last five estimates such that six estimates were used to calculate each weighted average." As previously recommended, EPA should provide information to enable readers to distinguish population-weighting used to develop a representative estimate from a source study using available estimates pertaining to particular subpopulations from procedures and calculations (adjustments) EPA used as part of a benefit transfer strategy.

4. The weighting process is more complex for the hedonic wage studies. For the Viscusi and Aldy (2007) study, VSL measures were constructed for each of 5 age groups. Although separate hedonic wage models were estimated for 1998, the weights appear to be for 2013 for the entire population. No adjustment was made to account for the difference between those who are working and those who are not for a variety of reasons. As a result, in this case the weights appear not only to be for the wrong year but the wrong population. This approach mixes a benefit transfer issue (assuming non-workers have the same VSL as workers) with the construction of a population mean based on a sample. The SAB has similar concerns about the EPA's weighting of the Aldy and Viscusi (2008) estimates and the weighting of any of the other hedonic wage estimates based on sub-populations. These observations provide examples where SAB is concerned that EPA distinguish the use of population weighting for the standard purpose of sample-weighting to derive a representative value from a particular study from explicit (or implicit) assumptions applied to implement a benefit transfer strategy or calculation.

Improving the Population Weighting Approach

In order to decide how great an effort should be placed on weighting subpopulations, the EPA could first determine whether there are large variations in VSL across subpopulations relative to variation across individuals. To improve the population weighting approach, future work could then be undertaken to investigate the possibility of developing a ~~more complex~~ set of subpopulation weights and benefit-transfer strategies that build upon what is known about the subpopulations covered in each of the available studies (whether currently included in analysis or not) in order to derive additional input to

Commented [K36]: Also what were the sampling weights from the orginal study for these sub groups

(observations for) the estimation of VSL. ~~This analysis~~An approach based on such studies c~~w~~ould eliminate the need for the screening criterion that studies necessarily provide a foundation, on their own, of a representative, population-weighted estimate of VSL. This approach could broaden the~~The analysis would mix standard benefit-transfer and statistically oriented benefit-transfer in a more comprehensive~~ foundation for estimation of VSL by ~~, potentially more consistent way and~~ enabling~~e~~ the use of a wider spectrum of available studies to derive VSL estimates for subpopulations. Meta-regressions over VSL estimates drawn from a larger set of studies, each of which might focus on subpopulations, could be conducted to develop a function that would allow adjustment for representativeness of the whole population. This approach could also be used to identify studies that appear to offer outliers in estimation, and then further consider whether there is reason to believe those studies may nonetheless offer valid insight to a portion of the distribution of values that may not be available from other studies. Such a meta-analysis would include statistical controls for methodological choices of the authors of studies.

An investigation of the feasibility of developing such an approach would involve consideration of the following questions:

1. Is it feasible to develop a weighting approach that builds upon multiple studies to improve estimation of VSL specific to many subpopulations of the U.S. and then aggregate such sub-population estimates to reach an improved, broadly representative estimate?

2. Would such a process be aided by including information from scientifically sound studies that focused on ~~a~~ narrower groups (e.g., ~~a~~ specific sub-populations), rather than setting the criteria for the included studies to arise from a broadly representative sample?

The SAB finds that such an approach could offer the advantage of including more information from more studies that may meet appropriate screening criteria while relaxing the requirement for a national focus at the level of the original studies drawn upon in support of a nationally representative population estimate. Although ~~a new and more complex~~ this approach could potentially reduce transparency, it could also improve confidence in a resulting representative estimate which would be based on a wider foundation of literature. This could raise confidence in benefit-cost analysis sufficiently to justify the loss of transparency ~~complexity~~ involved.

Commented [K37]: What is the theoretical basis for doing this need to connect back to maintained assumption about whether the population VSL varies by sub-group

Key Recommendations

- The White Paper should provide further explanation of how the weighting of subpopulation-specific estimates was actually done and how the studies were brought together for the aggregate estimate. The EPA should provide sufficient documentation to allow an independent party to replicate the analysis. This might be done in an appendix, perhaps with supporting information (e.g., a spreadsheet) that offers specific mathematical formulae used. In particular, the White Paper should:
  - Provide a more detailed explanation of how weighting procedures would affect estimates of standard errors.
  - More explicitly address the implications of selection bias.
  - Give much greater consideration to details of the specific studies being weighted. Appendix B mixes discussion of two kinds of procedures, population weighting and benefit transfer. It is important that population weighting be accomplished using standard procedures and that

benefit transfer assumptions be described and distinguished from weighting used to develop a representative estimate of VSL.

.

- Tie weights to the time period of the original study (at least for the development of a representative estimate supported by the original study) while aggregating available estimates across studies to obtain an overall estimate for 2013.
- Adjust for income differences in the populations (or time-periods) in individual studies after determining the estimates to be drawn from a particular study-time-period (income adjustments should then be addressed in the process of aggregating across studies to an estimate for a representative population).
- Explain how Hicksian and Marshallian measures of VSL should be aggregated with a consistent measure of income to account for income effects.
- Income adjustments of the VSL estimates derived from hedonic wage and contingent valuation studies must be consistent with the income concept relevant to each model, and in relation to the conduct of the original study and the timeframe of the source data. For hedonic wage models income is endogenous. However, for contingent valuation this is not the case and expected utility is being held constant. The analysis of proper treatment of income should reconcile these modeling assumptions before applying some adjustment.
- If income was used to form sampling weights then population weights to obtain a representative sample may involve income as a demographic variable to develop a representative estimate for a study population. This comment is not intended to suggest that deriving representative estimates necessarily use an income elasticity adjustment; rather, this comment concerns the statistical use of income as a demographic variable.

- The weighting process used for specific studies listed in Appendix B of the White Paper should be clarified.

- For future estimation of VSL, EPA should investigate the possibility of developing a set of subpopulation weights and benefit-transfer strategies that build upon what is known about the subpopulations covered in each of the available studies (whether currently included in analysis or not) in order to derive additional input to (observations for) the estimation of VSL. The intent would be to broaden the foundation for estimation of VSL by enabling the use of a wider spectrum of studies, although the gain in confidence achieved through use of a broader spectrum of studies could involve a reduction in transparency (while and should be mitigated to the extent possible by EPA would also providinge enough detail to enable a third party to replicate calculations).consider undertaking future work to investigate the possibility of developing a more complex set of subpopulation weights that build upon what is known about the subpopulations covered in each of the available studies (whether currently included in the White Paper or not).

Commented [K38]: Again this discussion is not consistent with the summary judgment

3.1.6. Estimation of Standard Errors

Charge Question 4. In some cases EPA estimated standard errors in the White Paper using information within studies or provided by the study authors, as described in Appendix B. Please comment on whether the methods used in the White Paper to estimate standard errors when such information was not readily available are appropriate and scientifically sound.

There are two major aspects of Charge Question 4 that must be addressed. One is related to how the standard error (se) of the VSL is calculated in situations when the standard error is not reported in the original study. In the White Paper, the EPA attempts to estimate the standard errors of the VSL when the original study does not report it. The SAB was asked to comment on whether the methods used to estimate the standard errors are appropriate and scientifically sound. The second, perhaps more important, aspect of the charge question is related to the methods the EPA used to estimate standard errors for the overall VSL estimates in the White Paper.

Calculation of the Standard Error of the VSL when it is Not Reported in the Original Study

The White Paper ~~fails to~~does not provide sufficiently detailed information about how the standard error (se) of the VSL is calculated in situations where one is not reported in the original study. The SAB recommends that the White Paper provide a detailed description of the method, including the formula ~~-~~ used for calculating the standard error for each study where the standard error of VSL is not reported. In particular, the SAB recommends that the EPA provide the following additional information in the White Paper.

1. For several stated preference studies, including Corso, Hammitt, and Graham (2001) and Viscusi, Huber, and Bell (2014), the White Paper states that standard errors for the reported VSL estimates were calculated using the confidence intervals reported by the authors. However, the White Paper does not provide details about how this was done. For example, u~~U~~nder some assumptions, one can calculate the standard error of a VSL estimates ($\widehat{VSL}$) based on its 95% confidence interval using the following formula:

$$se\left(\widehat{VSL}\right) = \frac{\overline{\widehat{VSL}} - \widehat{VSL}}{t_{0.025}(n)} \qquad (1)$$

   where $\overline{\widehat{VSL}}$ is the upper bound of the 95% confidence intervals~~ reported by the authors~~, and $t_{0.025}(n)$ is read off as the 2.5 percent point of the *t*-distribution with n degree of freedom. The White Paper should ~~present the formula~~describe the method it uses to translate confidence interval to standard error estimates.

2. For the Cameron, DeShazo, and Stiffler (2013) study, the White Paper states that "[w]e approximated the standard errors of the weighted VSL estimates the graphical information provided in an on-line appendix referenced in Figure 3 of the original study. We enlarged each graphic to visually identify an approximate point estimate for the 5th and 95th percentiles associated with each WTP estimate. We then used this information to calculate a standard error for each estimate." The SAB recommends that the EPA contact the authors to obtain the data instead of visually identifying an approximate point estimate for the 5th and 95th percentiles.

3. In several cases, the White Paper calculated standard errors for mean willingness to pay when the original study reported variance for median willingness to pay. The SAB recommends that in the White Paper the EPA provide a detailed explanation of how this was done.

4. For hedonic wage studies, the White Paper notes that the standard error of the VSL is calculated "based on the standard error of the risk coefficient alone." However, the exact formula used is

not provided. The SAB recommends that EPA include this information in the White Paper. If the study provides the average wage information, then there is sufficient information available to accurately calculate the standard error of the VSL. Specifically, assuming a log linear specification and that each worker works 50 weeks per year (i.e., treating this as a constant) for average wage $\bar{w}$, let $\hat{\beta}$ represent the estimated coefficient on the occupational fatality risk variable (i.e., the estimate of the true parameter β) and $se(\hat{\beta})$ its standard error. Assume risk is measured as the number of fatalities each year per 10,000 workers in the occupation-industry category. The estimated VSL is then given by (Aldy and Viscusi 2008):

$$\widehat{VSL} = 10{,}000(50)(\hat{\beta})(\bar{w}) = 50{,}000\hat{\beta}\bar{w}. \qquad (2)$$

This equation normalizes the VSL estimate to an annual basis by assumption of a 50-week work-year and by accounting for the units of the mortality risk variable. If the sample mean of wage provides an accurate estimate of the average wage $\bar{w}$, the standard error of the VSL is given by:

$$se(\widehat{VSL}) = 50{,}000\ \bar{w}\ se(\hat{\beta}).$$

On the other hand, if the sample mean of wage does not provide an accurate estimate of the average wage $\bar{w}$, and the original study treats the average wage estimate $\bar{w}$ as a random variable and provides a standard error estimate for $\bar{w}$,[1] the standard error of the VSL is given by:[2,]

$$se(\widehat{VSL}) = 50{,}000\sqrt{\left\{\left[se(\hat{\beta})^2 se(\bar{w})^2 + E(\hat{\beta})^2 se(\bar{w})^2 + E(\bar{w})^2 se(\hat{\beta})^2\right]\right\}}.$$

Methods for Estimating the Overall Standard Errors for the VSL

It is extremely important to provide accurate standard error estimates for the VSLs because standard errors are used to select the "preferred" model and the non-parametric estimate of the VSL. The SAB has reviewed the methods used in the White Paper to estimate the standard errors for the VSL estimates and provides the following findings.

1. Given the important role that standard errors play, the SAB finds that the White Paper ~~fails to~~does not provide detailed enough information about how standard errors of VSLs are estimated. ~~In fact,~~T~~ t~~here are ~~White Paper includes~~ only two short paragraphs to discuss the methods used to estimate the standard errors for the non-parametric VSL estimates (section 4.1.1). T~~ and does not discuss t~~he methods used to estimate the standard errors for the parametric VSL ~~estimates~~are not discussed at all.

2. The SAB finds that there are alternative, theoretically better, approaches (discussed below) to estimate standard errors for the overall VSL estimates.

[1] It should be noted that the sample standard deviation of wage might not provide a good estimate of how the mean wage estimate $\bar{w}$ deviates from the real average wage.

[2] The calculation assumes β and w are independent random variables makes use of the following formulas. The variance of the product of a constant a and a random variable X is given by $a^2 var(X)$. The variance of the product of two independent random variables X and Y is given by $var(X)var(Y) + var(X)[E(Y)]^2 + var(Y)[E(X)]^2$.

3. For the nonparametric approaches, the White Paper suggests five approaches/weighting methods for estimating the VSL. For each approach, the white paper uses a bootstrap method to estimate the standard errors of VSLs. The SAB finds that, because the discussion of the bootstrap methods is so brief, it is unclear how ~~the bootstrap approach~~this is implemented. For example, the paper states that "[t]o maintain the within-group correlation structure among the observations, we randomly drew *I* sets of groups with replacement from the primary sample of grouped observations. We did not re-sample observations below the top (group) level (Davison and Hinkley 1997 p 100-101, Ren *et al*. 2010)." (p. 25). It is not clear how each *I* set of groups was drawn and why ~~re-sample~~ observations below the top level were not re-sampled. ~~In fact, t~~The meaning of "group/data sample" is unclear. In footnote 11 on page 20, the White Paper states that "Hammitt and Graham (1999) and Corso, Hammitt, and Graham (2001) each examined ~~4~~ four samples." However, when looking at the last column of Table 6 on page 17, it appears that Hammitt and Graham (1999) examined only one sample and Corso, Hammitt, and Graham (2001) examined three samples. It is important to provide a clear definition of groups.

4. The White Paper uses a bootstrap approach to estimate standard errors for non-parametric VSL estimates. The SAB finds that there are conceptual ~~problems~~ issues with this approach. When the bootstrap approach is used, it seems that the estimated standard error reflects the variance of VSL estimates among the sample; it does not reflect the deviation of the VSL estimate from the true VSL. Conceptually, the accuracy of VSL estimates from individual studies used in the White Paper should affect the accuracy of the VSL estimates ~~in the White Paper~~. This means that standard errors of individual VSL estimates should affect the standard error of the overall VSL estimate. However, the bootstrap estimates of the standard error do not use the standard error estimates from the individual studies at all.

5. The SAB finds that there is an alternative, perhaps theoretically better way to calculate standard errors for each non-parametric VSL estimator. Specifically, by definition, the standard error of a non-parametric VSL estimate equals

$$se(\widehat{VSL}) = [E(\hat{y} - E\hat{y})^2]^{1/2} = \left[E\left(\sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}(y_{ij} - Ey_{ij})\}^2\right)^2\right]^{1/2}$$

$$= \left[\sum_{i=1}^{I}\sum_{j=1}^{m_i} w_{ij}^2(\sigma_\eta^2 + \sigma_\mu^2 + se_{ij}^2)\right]^{1/2} \qquad (3)$$

Thus, once σ_η^2, σ_μ^2, and se_{ij}^2 are estimated, one can use the above formula to estimate the standard error of the VSL estimate directly. The SAB recognizes the challenges in estimating σ_η^2 and σ_μ^2 , but the proposed approach has three main advantages: (1) it is based on theory, (2) it is consistent with the weighting strategy used, and (3) it uses the standard error estimates from individual studies.

6. The SAB finds that the White Paper does not currently provide any discussion about the approach used to estimate the standard error of VSL estimates for the hedonic wage approach. It

is assumed that the standard error of β_0 from the hedonic wage equation regression is used as the standard error of the VSL estimate. Again, it seems that the standard error of β_0 reflects the variance of VSL estimates among the sample; it does not reflect the deviation of the VSL estimate from the true VSL. Alternatively, because the hedonic wage regression provides estimates of σ_η^2 and σ_μ^2, one can calculate the standard error of the VSL estimate by using equation (3) above.

Key Recommendations

- ~~The white paper fails to provide detailed information about how the standard error of the VSL is calculated in situations where one is not reported in the original study. The SAB recommends that~~EPA should document precisely how the standard error of the VSL is estimated when the original study does not report one, so that an independent party could replicate the calculations. ~~the White Paper provide a detailed description of the method, including the formula used for calculating the standard error for each study where the standard error of VSL is not reported.~~

- The SAB ~~proposes~~ suggests an alternative, ~~perhaps t~~theoretically better, approach ~~way~~ to calculating ~~e~~ standard errors for both the ~~each~~ non-parametric and parametric- VSL ~~estimator~~estimates.

3.2. White Paper Analysis

3.2.1. Overall Methodology for Analyzing the Data

Charge Question 5. Please comment on whether the methodology used in the White Paper to analyze the data represents an appropriate and scientifically sound application of meta-analytic methods to derive generally applicable VSL estimates for environmental policy analysis.

The SAB finds that the meta-analytic methods used in the White Paper to analyze VSL estimates from the literature are, for the most part, ~~scientifically sound and~~ broadly consistent with standard and accepted practices for conducting meta-analyses. In some respects, it also offers advances beyond current common practices, for example in its decomposition of statistical error into three distinct components (group-level and observation-level non-sampling error and observation-level sampling error). To reinforce its consistency with common meta-analytic practices,~~this conclusion~~ it would be helpful for the White Paper to be more explicit about what these accepted practices are and how they are applied. This could be accomplished in several ways.

Commented [K39]: Not consistent with the statements above – **must be changed**

1. Several papers have proposed general steps, guidelines, and/or recommendations for conducting meta-analysis. The most relevant paper is Nelson and Kennedy (2009). This paper is referenced in the White Paper, but on a narrow~~er~~ issue. The White Paper would be strengthened by organizing the discussion around (or least referencing) these types of best practice guidelines. The White Paper does this to a limited extent with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework (Moher et al. 2009), but this really only applies to the study selection step.

2. The non-parametric statistical methods used in the EPA's analysis include approaches ("sampling error" and "total error" variance weighted mean) that are fundamentally similar to methods typically referred to in the meta-analysis literature as "fixed effect size (FES)" and "random effect size (RES)" methods.[3] Using, or least referring to, these labels, and describing how the methods used in the White Paper depart from these more standard practices, would help strengthen the presentation in the paper by tying it to the broader literature on meta-analysis. Also, when applying and comparing these nonparametric meta-analytic approaches, standard tests of homogeneity across groups (Q-tests) are generally recommended. These types of tests should be discussed and reported in the White Paper.

One of the principal best practice guidelines suggested by Nelson and Kennedy (2009) and supported by the SAB is to "ensure that the effect-size measures from the primary studies are all measuring the same thing." The White Paper could better address this recommended practice in several ways.

1. The White Paper should provide ~~more~~ detailed documentation about each of the primary studies and the selected value estimates in a way that would allow an independent party to replicate the results and that reinforces the direct comparability of the objects/commodities being valued. For example, it is ~~very~~ important that the temporal dimensions of the willingness to pay estimates be directly comparable (i.e., that they all measure or are converted to annual willingness to pay estimates for annual risk reductions). In the White Paper more attention should be given to describing the temporal features used in each study.

2. Where there are differences in the effect size measures across studies or value estimates, the White Paper should consider, discuss, and as appropriate include, adjustments to make the measures more comparable. For example, as previously discussed, the stated preference studies provide Hicksian value measures and the hedonic studies provide Marshallian measures. Also, whereas the stated preferences studies provide value estimates for non-marginal changes in risk, hedonic studies provide estimates of the marginal rate of substitution. The SAB recommends that EPA consider and describe the types of assumptions (e.g., preference structure) that would be needed to convert the Marshallian to Hicksian measures and the non-marginal to marginal values and evaluate the advantages or limitations of making these~~is~~ types of adjustments.

3. Although it is important to ensure that all included effect size estimates are measuring the same thing, t~~T~~he SAB finds that there is insufficient evidence in the income elasticity of VSL literature to adjust the VSL values from different studies to account for differences in income. Furthermore, gross domestic product per capita, which was used as the measure of income, has not increased for some income groups. Therefore, the SAB recommends that both the non-parametric and parametric analyses be conducted without this direct adjustment to VSL. ~~The parametric meta-regression analysis should include specifications with an income measure as an explanatory variable. This income measure should be selected to approximate as closely as possible the average disposable household income of the sample used in the primary study~~.

Another best practice guideline is to explicitly address and account for heterogeneity in the variance of the effect size estimates. The White Paper does this in several ways, including the use of "sample size

[3] The RES method is mentioned in the White Paper, but only in reference to the parametric/meta-regression approach.

weighted mean" in the non-parametric analysis. Sample size weighting has often been used in meta-analyses of willingness to pay estimates. Typically it is used as proxy for variance when variance estimates are not available. However, in this application variance estimates are available; therefore, it is not clear what is gained by including a sample size weighted approach. Its inclusion should be better justified.

Key Recommendations

- The White Paper should more explicitly discuss the standard and accepted practices for conducting meta-analysis, e.g., Nelson and Kennedy (2009), and how they have been applied. In particular:
  - The White Paper should refer to fixed effect size (FES)" and "random effect size (RES)" methods and describe how the methods used in the White Paper depart from these standard practices,
  - When applying and comparing nonparametric meta-analytic approaches, standard tests of homogeneity across groups (Q-tests) are generally recommended. These types of tests should be discussed and reported in the White Paper.

- The White Paper should provide more detail about each of the primary studies and the selected value estimates in a way that reinforces the direct comparability of the objects/commodities being valued.

- Where there are differences in the effect size measures across studies or value estimates, the White Paper should consider, discuss, and as appropriate include, adjustments to make the measures more comparable, and it should provide detailed documentation of any adjustments that are made so that an independent party could replicate the calculations.

- Both the non-parametric and parametric analyses should be conducted without adjusting the VSL values from the different studies to account for differences in income. There is insufficient evidence in the income elasticity of VSL literature to adjust VSL values from different studies to account for differences in income. Furthermore, gross domestic product per capita, which was used as the measure of income, has not increased for some income groups. ~~However, the parametric meta-regression analysis should include specifications with an income measure as an explanatory variable.~~

- The use of "sample size weighted mean" in the non-parametric analysis should be better justified to account for heterogeneity in the variance of effect size estimates.

3.2.2. Grouping Samples for Analysis

Charge Question 6. The White Paper classifies estimates into independent samples, also called groups, as described in Section 4. Estimates from some hedonic wage studies that use the same or very similar worker samples are grouped together for the analysis. Similarly, some of the stated preference estimates using the same sample are grouped together. Please comment on whether this methodology represents an appropriate and scientifically sound approach for accounting for potential correlation of results that rely on the same underlying data.

The SAB agrees that ~~that~~ it makes sense to group studies in the White Paper based on similar ~~samples~~ data sets to account for the lack of independence in estimates constructed from the samples. However, additional detail should be included in the White Paper to clarify how the grouping decisions were made. A column should be added to Table 6 of the White Paper to provide information more clearly identifying the composition of the various study groups.

The SAB also recommends that the EPA conduct additional analysis to check the robustness of the results to different plausible group definitions. Specifically, the SAB recommends that the EPA: (1) explore the sensitivity of results to alternative group assignments (e.g., grouping studies that used the same data set or the econometric approach together); (2) use the influence analysis to examine the robustness of the results to individually excluding each group; and identify; and (3) identify the primary estimate from each study and re-estimate the meta-regression using only primary estimates.

Key Recommendations

- Additional detail should be included in the White Paper to clarify how studies were grouped for the analysis. A column in Table 6 should provide information on the composition of various study groups.

- EPA should check the robustness of the results to different plausible study group definitions. This robustness check should include:
  - Exploring the sensitivity of results to alternative group assignments.
  - Using the influence analysis to examine the robustness of results to excluding each group.
  - Identifying the primary estimate from each study and re-estimating the meta-regression using only primary estimates.

3.2.3. Addressing Sampling and Non-Sampling Errors

> *Charge Question 7. Section 4.1 of the White Paper presents an expression that characterizes optimal weights that account for sampling and non-sampling errors, a framework that guides EPA's approach. Please comment on whether this is an appropriate and scientifically sound approach for addressing sampling and non-sampling errors.*

Additional information is needed to fully address this charge question. The SAB finds that derivation of the expression characterizing optimal weights that account for sampling and non-sampling errors should be more transparent in the White Paper. Therefore, the SAB recommends including in the text of the White Paper (or in an appendix) the various steps required to derive equation (4) in Section 4.1. Citations establishing the validity of the basic approach, ~~if not the literal~~ and the specific equation, should also be included. Regarding use of the weights, clarification of and justifications for the assumptions regarding the error components should be included. Finally, the white paper emphasizes the efficiency of the various estimators presented. The SAB ~~suggests~~ recommends that, without compromising best known, science-based practice for quantitative estimation, transparency should also be included as a criterion for selecting the estimator.

Key Recommendations

- The various steps required to derive equation (4) in Section 4.1 of the White Paper should be included in the text or in an appendix. Citations establishing the validity of the basic approach, if not the literal equation, should be included.

- With regard to the use of weights, clarification of and justifications for the assumptions concerning the error components should be included in the White Paper.

- Transparency should be included in the White Paper as a criterion for selecting an estimator.

3.2.4. Non-parametric and Parametric Approaches for Estimating Value of Statistical Life

> *Charge Question 8. The analysis in the White Paper adopts both non-parametric and parametric approaches (sections 4.1 and 4.2, respectively). Please comment on whether these approaches span a reasonable range of appropriate, scientifically sound, and defensible approaches to estimating a broadly applicable VSL for environmental policy and whether there are other methods that are more appropriate than those used in the White Paper.*

~~The EPA White Paper adopts both non-parametric and parametric approaches to estimate a VSL.~~ The SAB finds that some additional information is needed in the White Paper, especially to explain the use of the nonparametric approach ~~to explain the use of these approaches~~. The SAB recommends that the EPA provide citations for the non-parametric approaches (estimators 1-5 on pages 22-23 of the White Paper) and better justification for the methods in terms of the specific application. Specifically, the justification should explain why these methods are relevant to finding the central tendency of VSL estimates from studies that in most cases report multiple estimates. Some discussion of the conceptual merits and data requirements of each method is needed. Calculations should be documented with sufficient detail to allow a reader to know precisely how to replicate the calculations. The SAB notes that estimator 3 is described in the text on meta-analysis by Hunter and Schmidt (2004) and estimator 4 is described in the text on meta-analysis by Hedges and Olkin (1985) and implemented in a recent meta-analysis by Hsiang et al. (2013).

The SAB agrees with the EPA's conclusion that the mean of group means estimator is the preferred non-parametric method. The EPA's argument is that it has the smallest estimated standard error (p. 32 of the White Paper). The SAB recommends that the EPA emphasize the rationale that ~~also justify~~ use of the mean of group means estimator ~~on the grounds that it~~ avoids giving too much ~~more~~ weight to studies~~papers~~ that report multiple~~more~~ estimates. It is not clear why there is so much variation across papers in the number of reported estimates. This may be a result of idiosyncratic factors (e.g., stylistic choices by authors, requests by referees for robustness checks) and, as such, it is better to give equal weight to groups of estimates.

Although the SAB supports the EPA's focus on the mean of group means estimator, it is recommended that the agency explore the use of an alternative non-parametric method that incorporates information on sampling error variance from each study. This estimator is a blend of approaches 2 and 4 and would be computed as follows:

$$\hat{y} = \frac{1}{I}\sum_{i=1}^{I}\frac{1}{\sum_{j=1}^{m_i} se_{ij}^{-2}}\sum_{j=1}^{m_i} se_{ij}^{-2} y_{ij} \qquad (4)$$

The estimator computes the mean of sampling error variance weighted group means.

For the parametric estimator, the SAB recommends that the EPA provide better explanation of and justification for the included control variables. Some of this discussion is found in section 6.1 of the White Paper, but is better placed in section 4.2. The SAB recommends that, if feasible, the EPA should include additional controls in the parametric model. One suggestion is to include indicator variables for a study having specific major contributors to the VSL literature as co-authors.~~dummy variables for whether given researchers are co-authors.~~

The parametric model used to estimate the VSL included a time trend variable for the year the data were collected for the study from which the primary estimate was drawn (see equation 16 in the White Paper). The SAB recommends that time trend variables not be included in either the parametric or the non-parametric models. The SAB concludes there is not a rationale for giving different weights to estimates from different years and, thus, recommends the use of equal weights in forming the average VSL. However, to explore whether older or new studies have a strong influence on the VSL estimate, the EPA should consider conducting a sensitivity analysis similar to the influence analysis in Table 10 of the White Paper.

~~There is evidence of an upward time trend in inflation-adjusted VSL estimates, whether or not the estimates are further adjusted for income differences over time. The parametric model includes a control for this time trend but the non-parametric estimators does not. The current document does not explain what the time variable in the parametric model was intended to capture. It could reflect real changes in preferences over time, changes in knowledge about health risks, or changes in valuation methodology. The SAB recommends that the EPA be consistent in its treatment of the time trend. If it is controlled for in the parametric model, it should be controlled for in the non-parametric models. However, it is not clear how the time trend, if included, should be used in producing the final VSL estimate. The time variable could be evaluated in the current year, the end of the data period, or the midpoint of the data period. The first two approaches would place greater weight on more recent studies, which could be appropriate if one believes that methodologies are improving over time. However, if one believes that methodologies are getting better, estimates should simply be taken from the most recent study. More generally, if one believes there are differences in quality over time, one should pick the study that appears to be the best. Leaving the time trend out of the models would implicitly give equal weight to estimates from different years. This approach might be preferred in the absence of strong beliefs about the source of the time trend.~~

Key Recommendations

- Citations should be provided in the White Paper for the non-parametric approaches (estimators 1-5 on pages 22-23) and better justification should be provided for the methods in terms of the specific application. Specifically the justification should explain why these methods are relevant to finding the central tendency of VSL estimates from studies that in most cases report multiple estimates.

- EPA should justify use of the mean of group means estimator on the grounds that it~~avoids giving~~ avoids giving too much weight to studies that report multiple estimates. ~~more weight to papers that report more estimates.~~

- EPA should explore the use of an alternative non-parametric method that incorporates information on sampling error variance from each study.

- For the parametric estimator, the SAB recommends that the EPA provide better explanation of and justification for the included control variables.

- EPA should not include a time trend variable in either the parametric or non-parametric models, but should consider a sensitivity analysis to determine whether older or newer studies have a strong influence on the average VSL.~~be consistent in its treatment of the time trend time trend in VSL estimates. If it is controlled for in the parametric model, it should be controlled for in the non-parametric models.~~

3.3. White Paper Results

3.3.1. Proposed Estimates of Value of Statistical Life

Charge Question 9. The White Paper presents estimates using parametric and non-parametric models, pooled across stated preference and hedonic wage studies as well as balanced (i.e., equal weight to each study type), and weighted using different approaches. Of the range of estimates presented (see Section 4) the White Paper proposes the use of estimates from the following models:

- *Non-parametric model, balanced, mean of study mean*
- *Parametric, balanced*

Please comment on whether these proposed estimates represent reasonable and scientifically sound conclusions from the analyses in the White Paper and whether there is a different set (or sets) of results that are preferable based on the data and analysis in the White Paper.

The SAB finds that the proposed summary VSL measures (nonparametric balanced mean of the study mean and parametric balanced) are defensible estimates. As indicated in the response to Charge Question 8, the EPA should also consider using the non-parametric sampling error variance weighted group mean in place of the non-parametric mean of group means estimator. However, it is important to distinguish these estimates from those that have used weights to construct "general population" measures for the U.S. population. There are inconsistencies in the weights used across studies. For the estimates derived from the 2002 stated preference study designed by Cameron and DeShazo for a representative sample of U.S. households, the weights should be based on the Knowledge Network weights for the 2000 census not the 2010 census. For the hedonic wage studies, the weights appear to be based on 2013 information for the general population when the samples are for earlier years and are designed to represent populations of individuals who choose to work full time. In this case, weighting to derive a mean for the general population mixes a *benefit transfer* assumption with a *sample weighting* decision. The benefit transfer assumption involves assuming non-workers whether unemployed, retired or not participating for another reason have the same risk tradeoffs (VSL) as those working. This decision should be explicitly stated and not simply subsumed in~~explicit and not "buried" in~~ the weighting process. Although this may not cause a large effect, it is not possible to determine the size of the effect based on what is presented.

Commented [K40]: NO must be changed –not defensible –this is inconsistent with everything here!!

Adjustment of VSL estimates by an income elasticity of VSL and index of income growth (based on GDP per capita) does not seem appropriate. However, conversion of VSL to inflation adjusted dollars

would be appropriate. "Building in" the income elasticity and growth assumptions as maintained hypotheses before constructing the mean mixes a benefit transfer decision with an adjustment for household income across different studies. More specifically, income adjustment could involve: (1) adjustment for differences in the income across different samples that could hypothetically alter the risk tradeoff; and (2) adjustment for changes in real income over the time period covered by the effects of a rule where assumptions about the growth of the income might be expected to raise all households income for the future date when the policy was implemented. This type of "income adjustment" would be a part of the benefits transfer associated with modifying a unit value so it is consistent with the economic conditions at the time the policy is assumed to affect mortality risks. It is different from an adjustment conducted to a primary estimate before developing the mean estimates for VSL. At present, the documentation of income adjustment in the white paper is not clear. Table 6 of the White Paper refers to the use of an income elasticity of 0.7 but does not clearly discuss the income used in the two adjustments. In addition the SAB notes that adjustment for income with the stated preference measures would need to be different because these are derived from Hicksian welfare measures (Smith et al. 2002; 2003)~~). *<<Chair's note: can we provide a citation for methods that could be used for this kind of adjustment?>>.*~~

Key Recommendations

- The proposed summary VSL measures (nonparametric balanced mean of the study mean and parametric balanced) are defensible estimates but as indicated in the response to Charge Question 8, the EPA should ~~also consider~~ explore the use of an alternative non-parametric method that incorporates information on sampling error variance from each study.~~using the non-parametric sampling error variance weighted group mean in place of the non-parametric mean of group means estimator.~~

- The documentation of income adjustment to VSL should be clarified in the White Paper. Adjustment of VSL estimates by both an income elasticity of VSL and index of income growth (based on GDP per capita) does not seem appropriate. However, conversion of VSL to inflation adjusted dollars ~~would be~~is appropriate.

Commented [K41]: No they are not –we cannot judge

Commented [K42]: Is not appropriate –not does not seem appropriate

3.3.2. Influence Analysis

Charge Question 10. The results section of the White Paper concludes with an influence analysis. Please comment on whether this analysis is a reasonable way to characterize the influence of individual studies on the estimated VSLs, whether the results of the influence analysis suggest any changes or modifications to the estimation approach, and whether it is important to include an influence analysis.

An influence analysis is important, especially given the implicit assumptions underlying the structure of the non-sampling error related to groups and given the relatively small number of VSL estimates. Some form of influence analysis is important for meta-analysis in cases where there are few studies to consider, and therefore one or two individual studies might have a substantial influence on the estimates. Influence analysis is most important to make sure that the influence not skew the results in a single direction. For example, if there are two studies with +10% and -10% influence the two studies are more

or less balanced. ~~Looking at~~ With regard to the mean of group means in the White Paper, the two most influential studies are Corso Hammitt and Graham (2001) at -13.8% and Chestnut, Rowe, and Breffle (2012) at +~~-~~11.1. Taken together, these studies nearly balance each other. In contrast, for the maximum likelihood stated preference estimates, the Corso, Hammitt and Graham (2001) at -22.8 % is well over two times more influential than the second most influential study, which fortunately is of the opposite sign. Rather than dropping Corso, Hammitt and Graham (2001) altogether, one might use a robust estimation techniqu~~e~~ that limits the influence of this observation. One possibility is to adjust the weight on this study downward until it just balances the Alberini et al. (2004) study, or to down weight all studies that are identified as relatively influential (perhaps studies that fall above the +/- 10% influence range). This type of approach of down weighting highly influential observations has a long history.

It would also be useful to consider the potential for using regression diagnostic indexes (Belsley et al. 1980; Cook and Weisberg 1982; Belsley 1991) for the parametric modeling of VSL. These statistics allow analysts to consider whether specific observations were influential to individual coefficients in the meta-regression function. They allow an assessment of whether the magnitude and significance of individual coefficients was influenced by particular observations.[4] Since these correspond to the specific studies and models within a study, they could help in understanding how the group definition discussed earlier influences the specific mean statistics proposed to construct a population level measure for the mean VSL.

Key Recommendations

- Influence analysis of the maximum likelihood stated preference estimates indicates that Corso, Hammitt and Graham (2001) at -22.8 is well over two times more influential than the second most influential study. The EPA should consider using a robust estimation technique that limits the influence of this observation.

- The EPA should consider the potential for using regression diagnostic indexes (Belsley et al. 1980; Cook and Weisberg 1982; Belsley 1991) for the parametric modeling of VSL.

3.4. Protocol for Future Revisions of Value of Statistical Life

3.4.1. Criteria for Inclusion and Exclusion of VSL Estimates in Future Analyses

> *Charge Question 11. In the previous SAB advisory report (U.S. EPA Science Advisory Board 2011), the SAB endorsed the idea of establishing a standardized protocol and regular schedule for future updates to the Agency's mortality risk valuation estimates. Please comment on relevant statistical criteria for the inclusion of additional eligible estimates and/or the exclusion of older estimates that could help inform the development of a standardized protocol for future updates and the timing or frequency of those updates.*

[4] These are "old" references but can provide useful indexes of how specific observations influence results. The discussion of "short data" in chapter 7 in Belsley (1991) may be especially relevant to parametric models developing meta summaries with limited variation in the risk and/or income measures that are used to estimate income elasticities or scope effects.

The SAB provides general and specific recommendations on the development of a standardized protocol for future updates and the timing or frequency of those updates.

General Recommendations

The SAB notes that the value of risk reduction for mortality (VRR), a term previously suggested by the SAB as a replacement for VSL, is very likely the most important "benefit measure" used in EPA's benefit-cost analyses for policies related to mortality risk. The level of staff effort and other research resources devoted to regularly updating and refining VRR estimates should be commensurate with their importance for policy evaluation.

Given the importance of VRR, high priority should be assigned to increasing the pool of high quality studies to support the VRR meta-analysis. This is particularly important due to the small number of data sets to support hedonic price estimates, and the relatively small number of stated preference studies currently included in the meta-analysis. In addition to improving the precision of VRR estimates, additional high quality studies could improve the ability to estimate other important characteristics of VRR, such as possible time trends, income elasticities, variability over subpopulations, cancer premiums, and other factors that are central to policy analysis.

In the immediate term, the EPA can expand the number of studies by considering whether useful information can be extracted from a variety of studies previously excluded from VRR calculations. Subject to caveats and recommendations detailed below, such studies might include those with samples that are not representative of the national population, results from other economic studies of risk preferences (e.g., transportation safety, risk-risk tradeoffs, etc.) and results from research outside of traditional peer-reviewed journals. The SAB could provide guidance on which studies are appropriate, and/or any additional steps necessary to ensure that studies are used in an appropriate manner. Recommendations regarding these issues are detailed below.

In the long term, new high quality studies could be elicited by EPA using existing and new mechanisms. For example, EPA should consider whether estimation of VRR and its various attributes (e.g., time trends, etc.) should be a high priority topic for Science to Achieve Results (STAR) grants and fellowships, EPA sponsored conferences, special issues of journals, and young researcher awards. The EPA could even consider the feasibility of sponsoring its own refereed journal that focuses on analyses of direct relevance to meeting the agency's needs.

The EPA might also obtain more general information about protocols for updating estimates from the experience of other agencies that construct economic index numbers for policy. For example, the Bureau of Labor Statistics (BLS) calculates the Consumer Price Index, based on a weighted set of prices to calculate the cost of a "representative basket" of consumer goods and services. BLS updates the weights periodically to account for new goods and for changes in quality of goods over time. The EPA could learn from protocols used by BLS and other agencies for periodic updates.

Statistical Criteria for the Inclusion of Additional Eligible Estimates and/or the Exclusion of Older Estimates

There should be a single set of criteria for determining which studies are of sufficient quality to be included in the estimates of VRR. Therefore, the validity criteria for inclusion of additional studies and exclusion of older studies should be the same as those to assess the estimates that are currently in use. These criteria have been discussed in detail in the response to Charge Question 1a.

Similarly, the SAB recommends that the exclusion of older estimates be evaluated on a case-by-case basis using the same validity criteria, rather than dropping studies simply based on their being dated, *per se*. If there is strong evidence that risk preferences change over time, the SAB recommends developing procedures to adjust older estimates that are otherwise judged to be valid, rather than dropping estimates simply because they are older. This is especially pertinent given the small number of studies upon which current VRR estimates are based.

Timing or Frequency of Updates

The SAB finds that a 5-year interval of updating estimates is probably appropriate. More frequent updating might be desirable, but based on experience in the past several years, there appear to be too few new estimates each year to justify the time and expense involved in more frequent updating.

Potential Sources of Information Outside of Peer-Reviewed Journals ***<<Chair's note: this should be discussed>>***

As previously indicated, increasing the number of high quality studies included in the meta-analysis is a high priority. For this reason, the SAB has considered whether studies should be restricted to those that are published in peer-reviewed literature.

The SAB recommends that the EPA not necessarily restrict studies to those published in peer-reviewed journals, but rather that studies outside of the peer-reviewed journals should be considered for inclusion following a transparent and rigorous peer review process. The SAB emphasizes that it is inadequate to simply assert that a study was subject to peer review. Rather, a quality controlled peer review process should be established. For example, EPA might ask the SAB to organize a process to review research results outside of traditional peer reviewed journals, both to identify appropriate reviewers (possibly including SAB members), and to determine whether or not studies that undergo peer review are judged to "pass" the review process, and therefore qualify for inclusion.

Extending sources of information to research outside of peer-reviewed journals has the potential to substantially increase the number of studies available to estimate VRR, and research papers outside of peer-reviewed journals likely include high quality empirical analyses, even if they are not submitted for publication in journals. A major challenge to relying only on publications in peer reviewed journals is that economics journals rarely publish articles that contain routine empirical analyses without some sort of innovation or other improvement in the state-of-the-art in economic theory or empirical methodology. In contrast to some other disciplines, the field of economics places a low priority on improvements in the state-of-the-inventory of empirical knowledge. This severely discourages production of studies serving a primary function of recording value estimates useful for policy analysis. As a consequence,

many analyses could provide satisfactory estimates of VRR, but may not be submitted to peer-reviewed journals, or may be rejected for publication because they do not improve upon the state-of-the-art in economic theory or empirical methodology. This may be particularly relevant for analyses carried out by consulting companies for government agencies, for whom publication of research results in peer-reviewed journals may or may not be of high priority.

Information from Other Economic Studies of Risks

The SAB recommends that the EPA consider whether useful information can be extracted from other studies that could improve estimates of VRR and its characteristics (e.g., latency, morbidity). This might include studies of risk-risk tradeoffs, hedonic analyses in addition to hedonic wage studies, risk studies in the transportation safety literature, and possibly others. For example, EPA might consider using ~~the~~ results of ~~a~~ risk-risk studies~~y~~ that employ~~ed~~ a stated preference approach, wherein respondents were asked to choose whether to undergo treatment (e.g., a risky surgery) that has a stated risk of immediate mortality versus a given risk of cancer, which involves stated risks of both long term morbidity and subsequent mortality. EPA might also use the results of a study that asked respondents to choose whether to undergo treatment that has a stated risk of morbidity (e.g., paralysis, chronic pain, etc.) versus foregoing treatment, in which case they face a stated mortality risk (Hauber et al. 2013). These studies could potentially be useful for calibrating differences in VRR across risks with differing degrees of latency, morbidity, etc. (e.g., a possible cancer premium). These issues are particularly relevant given EPA's focus on environmental risks, which often involve long latency periods, and where mortality is often preceded by a significant period of morbidity.

The EPA should also consider whether useful information can be extracted from other categories of studies, such as hedonic literature outside of hedonic wage studies, the literature on health care cost effectiveness analysis, and possibly transportation safety studies. For example, it may be possible to extract useful VRR information from hedonic studies of the effects of air pollution on housing prices, although challenges may exist in isolating mortality and morbidity effects from other effects, such as visibility. Studies on health care cost effectiveness analysis use measures of health-related quality of life that often fall short of utility-theoretic standards but could nevertheless be useful. There is a comprehensive searchable database of such studies that is managed by Tufts University (Tufts University Medical Center, 2016). Estimates of VRR from transportation safety studies might also be applicable VRR from environmental risks.

Information from Studies with Non-National Samples

Similarly, the SAB recommends that EPA not necessarily exclude studies simply because they are based on non-national samples, as long as there is a broad set of studies that as a group is generally representative of the nation as a whole or can be used to either develop a representative estimate for the nation as a whole, or improve the representation of VSL values of subpopulations that are underrepresented or omitted from studies used to estimate a representative value for the nation as a whole . For example, EPA should consider studies based on representative samples at the state and regional levels, as long as there is an adequate number of studies using representative samples for a diverse set of states and/or regions. The SAB suggests that it probably would not be appropriate to adopt estimates from studies based on narrow demographics, or a very small geographic area (e.g., a single community) since they may not be representative.

If there are a reasonable number of studies at the state and regional levels, one could carry out consistency checks to ensure that similar estimates result from national level studies and a set of state and/or regional level studies. As previously indicated, in addition to improving the precision of VRR estimates, increasing the number of high quality studies has the advantage of allowing improvements of estimates of related measures, such as time trends, income elasticities, and variability over subpopulations.

Open Data Initiatives

Another challenge in depending only on existing studies published in peer-reviewed journals for VRR estimates is it that makes EPA dependent upon only those results that are reported in the publication, and possibly additional information that can be obtained by contacting the authors. For example, some studies report VRR estimates, but do not report associated standard errors or confidence intervals on VRR, income elasticities, and estimates by sub-population. Additionally, different studies use different statistical methods, control for different influences and otherwise use different procedures that are difficult to control for after the fact.

The SAB recommends that the EPA consider a long term strategy of requiring that a more inclusive set of research results, and even whole data sets, be made generally available for use by the research community and by government agencies. Project Open Data (U.S. Office of Management and Budget and U.S. Office of Science and Technology Policy 2016) provides an excellent framework for making data available in order to improve the information obtained from available studies.

It is becoming increasingly common practice for agencies and professional associations to develop open data policies, which require or strongly encourage that data be made widely available to the research community, to the extent feasible. For example, in May 2013 President Barack Obama issued an Executive Order and an associated Open Data Policy for all federal agencies. The Office of Management and Budget's Open Government Directive creates a "presumption in favor of openness to the extent permitted by law and subject to privacy, confidentiality, security, or other valid restrictions," and requires that agencies publish high-value data sets in an open format through Data.gov (The White House 2016). ~~Federal grants and contracts could require that data collected under~~ EPA could also require ~~grants and contracts require~~ that data collected under grants and contracts awarded by the agency be published to Data.gov in standard format (U.S. General Services Administration 2016), unless there is a compelling reason that the data not be published. Such a policy might allow exceptions and be subject to possible censuring of individual variables and observations as necessary to ensure protection of confidentiality. This is consistent with U.S. Office of Management and Budget (OMB) policy, which established the principle that, where feasible, data be public, accessible, fully described, reusable, complete, timely and managed post-release. Similar open data policies have been adopted by peer-reviewed journals like *Science*, *Nature* and *PLOS*.

An open data policy would have the advantage of providing an opportunity to replicate research results, to help improve quality control on reported estimates, and to go back after the fact to estimate parameters of importance that are not reported in the original publication (e.g., VRR standard errors). Additionally, data from multiple studies could be used to apply more refined estimation techniques, to apply more comparable standards (e.g., explanatory variables) across studies, and to correct possible

biases in studies. For example, data collected in the immediate aftermath of a major event (e.g., the Great Recession of 2007-2009) might not be representative of the long term. A single parameter estimate from a study using pooled data from 2005-2010 might not be refined enough to adjust for differences during the recession years. Access to the original data set could provide researchers with the opportunity to ~~correct for~~adequately take into account such influences.

More broadly, collecting primary data is expensive, and it is inefficient to expend large amounts of funding to collect data for a single analysis and then exclude those data from being used for other productive purposes. Indeed, a recent report has estimated that open data could add $3 trillion to $5 trillion in economic value to the global economy each year (Manyika et al. 2013). While the SAB has not had an opportunity to review this particular study, it is clearly suggestive of the substantial social value in making data more widely available to the research community.

At the same time, the SAB recognizes there are important challenges to making data sets publicly available. For example, issues may arise with respect to confidentiality of survey respondents in some data sets. Also, all data sets have important limitations that are often best known to those who originally collected that data. In addition, many researchers will want to publish results from data sets prior to making them public. However, the SAB finds that challenges associated with these issues can be minimized by carefully considering data sharing policies and the important efficiencies in making data publicly available. The SAB also notes that making data publicly available after a reasonable amount of time would fit into the process of updating the VSL estimate. The SAB recommends that the EPA work in collaboration with other agencies and professional associations to pursue reasonable and prudent actions to make data publically available. For example, the EPA could learn from the policies established by the National Science Foundation program for Long Term Ecological Research.

Commented [K43]: This is now controversial at least with respect to the LTER working on social sciences

Routine Compilation of Existing Data Sets

The EPA might also make an effort to routinely compile data from various key sources for regular use. For example, as previously indicated, the EPA might simplify periodic updating of hedonic wage estimates of VRR by creating an archive of wage data and perhaps other data from the U.S. Census Bureau's demographic supplement to the Current Population Survey, matched with data from the U.S. Bureau of Labor Statistics Census of Fatal Occupational Injuries (CFOI) in standardized form, and perhaps other data sets. Once in place, such a data archive would allow for consistent periodic updates of VRR at low cost, rather than waiting for updated publications in the peer-reviewed literature. This approach also has the advantage of providing a consistent methodology underlying hedonic wage estimates over time. EPA might create its own data archive, or the compiled data might be published in existing data archives, such as Data.gov.

Key Recommendations

- The pool of high quality studies to support the VRR meta-analysis should be increased. To accomplish this the EPA should:
  - Consider whether estimation of VRR and its various attributes (e.g., time trends, etc.) should be a high priority topic for Science to Achieve Results (STAR) grants and fellowships, EPA sponsored conferences, special issues of journals, and young researcher awards

  - Consider the feasibility of sponsoring its own refereed journal that focuses on analyses of direct relevance to meeting the agency's needs.
  - Obtain more general information about protocols for updating estimates from the experience of other agencies that construct economic index numbers for policy.

- There should be a single set of criteria for determining which studies are of sufficient quality to be included in current and future estimates of VRR.

- A 5-year interval for updating VRR estimates is appropriate.

- EPA should not restrict studies used for updating VRR to those published in peer-reviewed journals. Studies outside of the peer-reviewed journals should be considered for inclusion following a transparent and rigorous peer review process.

- The EPA should consider whether useful information can be extracted from other studies that could improve ~~estimates of~~ understanding of VRR estimates and how they relate to underlying characteristics ~~and its characteristics~~ (e.g., latency, morbidity). This might include studies of risk-risk tradeoffs, hedonic analyses in addition to hedonic wage studies, and risk studies in the transportation safety literature.

- The EPA should not exclude studies based on non-national samples from use in updating VRR as long as there is a set of studies that as a group is representative of the nation as a whole.

- EPA should consider a long term strategy of requiring that a more inclusive set of research results, and even whole data sets, be made generally available for use by the research community and by government agencies.

3.4.2. Valuing Reductions in Risks of Cancer

Charge Question 12. In its 2011 report the SAB-EEAC recommended "…EPA work toward developing a set of estimates…for policy-relevant cases characterized by risk…" (U.S. EPA Science Advisory Board 2011, pp. 10). Among the studies that meet the selection criteria in the current White Paper, three stated preference studies provide values for reductions in risks of cancer (i.e., Hammitt and Haninger 2010, Chestnut, Rowe, and Breffle 2012, and Viscusi, Huber and Bell 2014). Only two of those studies (Hammitt and Haninger 2010 and Chestnut, Rowe, and Breffle 2012) allow for a within study comparison of values for cancer and non-cancer risk reductions. However, EPA could augment the literature by modifying the selection criteria to include studies from other countries or from the grey literature, and/or using other methods (e.g., risk-risk studies). Please comment on whether, and if so how, selection criteria for identifying studies for estimating a cancer differential should differ from those used in the current White Paper. Does the literature support a non-zero cancer differential?

The SAB has previously concluded that "research suggests that people are willing to pay more for mortality risk reductions that involve cancer than for risk reductions from accidental injury and proposes a placeholder value that could be used for this cancer differential while the Agency pursues long-term research to differentially value other types of risks" (U.S. EPA SAB 2011). The motivation behind a

potential cancer differential is that a death from cancer is preceded by a significant period of morbidity, while a death from accidental injury may not be[5]. According to this motivation, a cancer death can be thought of as two events, a period of morbidity followed by an early death. Moreover, cancer treatment typically is accompanied by surgery, chemotherapy and radiation that can have serious debilitating side effects. The experience of death is also traumatic for family and friends as well as the affected individual in ways that sudden death is not. Logically, one would expect a premium to be associated with reduction of risk of mortality that also includes longer and/or more severe periods of morbidity. ~~Logically, a~~A death preceded by a significant period of morbidity would be viewed as worse than a sudden accidental death (though there may be some benefit to being given a period of time to put one's affairs in order). Indeed, Gentry and Viscusi (2016~~5~~), using revealed preference wage data, find that wage premiums for occupational mortality risks that tend to be preceded by longer periods of morbidity are higher than premiums for occupational mortality risks that tend to be preceded by shorter periods of morbidity, and that the value of a statistical life can be decomposed into a value of the fatality risk plus the value of the associated morbidity risk. These studies show that people value both mortality risks and associated morbidity risks, suggesting that a cancer premium could exist.

Given that a cancer premium is possible, is there enough evidence in the literature to establish its size? Few studies have done clean comparisons of an estimated VSL for cancer-related deaths to a VSL for sudden death. Hammitt and Haninger (2010) found that willingness to pay to reduce risk of death from disease caused by consumption of pesticides was larger, but not statistically different from willingness to pay to reduce risk of death from an automobile accident. Chestnut, Rose, and Breffle (2012) found that willingness to pay to reduce risk of death from cancer was larger, but was not statistically different from willingness to pay to reduce risk of death from heart attack. Cameron and Deshazo (2009) compared VSL for sudden death to VSL for an illness profile that involved one or five years of illness followed by death. They found that willingness to pay for a risk reduction was not significantly different across these three treatments, though this comparison confounds morbidity and latency.

One study that did claim to find a cancer differential was Viscusi, Huber and Bell (2014). They estimate a VSL for a cancer death of 10.85 million dollars. They compare this VSL to the median value of the VSL for an accidental death estimated from several studies, which they find to be 9 million dollars. From this they conclude that there is a positive cancer differential of twenty-one percent. Several points should be made about their findings. First, the 10.85 million dollar VSL estimate is based on a VSL of 8.1 million dollars for a cancer death with a ten year latency. The 10.85 million dollar value was arrived at by discounting over ten years at a discount rate of three percent. People may use some other method than financial discounting to tradeoff between current health risks and future health risks. Second, while Viscusi, Huber and Bell (2014) present confidence intervals for their VSL estimates, their own robustness checks show that the estimated VSL for cancer risks is sensitive to their analytical approach. Viscusi, Huber and Bell (2014) elicited willingness to pay values using a multiple-bounded dichotomous choice method. They found, as is often the case, that the estimated VSL differs depending on whether only the first response is used in the analysis or all responses are used. Specifically, they found that using all three responses per respondent resulted in a VSL estimate that was thirty-one percent higher than the VSL estimate based on only the first response. Had they used only the first response in their analysis, they would have concluded that the value of a cancer VSL was actually less than the median

[5] It should be noted that a second motivation for a positive cancer differential has been proposed, namely that people associate a higher level of dread with cancer risks than with other health risks (Sunstein 1997, for example).

VSL value for accidental deaths. Based on available studies, the SAB concludes that there is not sufficient evidence at this time to justify a non-zero cancer differential.

The SAB recommends that, instead of adopting a nonzero cancer differential, the EPA consider using existing methods to value the morbidity that occurs prior to an early death, and add that estimated morbidity value to conventional estimates of the value of the associated mortality. The EPA currently values morbidity from cancer in cases where the cancer is not fatal, but does not value morbidity in fatal cancer cases. The EPA should value cancer morbidity regardless of whether that morbidity leads to an early death. This recommendation also applies to other environment-related mortality risks, including cardio-pulmonary disease. Morbidity that occurs prior to an early death should be valued in all cases. Mortality can then be valued using conventional VSL estimates.

The EPA should encourage and support ongoing research on whether willingness to pay to reduce the risk of an early death preceded by a period of morbidity is correctly valued by summing the value of the morbidity plus the value of the mortality. At this time, the SAB does not have evidence to suggest that that approach would over- or under-state the true willingness to pay. Gray literature studies, studies conducted outside the United States, and studies that do not directly estimate VSL, such as risk-risk tradeoff studies and risk-benefit studies (see studies cited in Appendix C), could be assessed to determine whether there is evidence that the VSL for different mortality risks differs, after having controlled for the value of associated morbidity. <<***Chair's note: can we provide citations to relevant studies?***>> Such studies can help the EPA and the SAB determine whether the SAB's recommendation should be reassessed. However, if and when it is determined that a cancer differential (or a differential for other diseases) is justified, the same selection criteria should be used to identify studies to measure the differential(s) as is used to identify studies to establish the baseline VSL.

Key Recommendations

- Based on available studies, the SAB concludes that there is not sufficient evidence at this time to justify a non-zero cancer differential. The SAB recommends that, instead of adopting a nonzero cancer differential, the EPA consider using existing methods to value the morbidity that occurs prior to an early death, and add that estimated morbidity value to conventional estimates of the value of the associated mortality.

- The EPA currently values morbidity from cancer in cases where the cancer is not fatal, but does not value morbidity in fatal cancer cases. The EPA should value cancer morbidity regardless of whether that morbidity leads to an early death. This recommendation also applies to other environment-related mortality risks, including cardio-pulmonary disease.

- The EPA should encourage and support ongoing research on whether willingness to pay to reduce the risk of an early death preceded by a period of morbidity is correctly valued by summing the value of the morbidity plus the value of the mortality. At this time, the SAB does not have evidence to suggest that approach would over- or under-state the true willingness to pay.

3.5. Income Elasticity of the Value of Statistical Life

3.5.1. Income Elasticity Literature

> *Charge Question 13. The EPA document Technical Memorandum: Income Elasticity presents a summary of the recent income elasticity literature based on a review presented in Robinson and Hammitt (2015). Please comment on whether Robinson and Hammitt (2015) and the EPA Technical Memorandum provide an appropriate and scientifically sound summary of the income elasticity of VSL (IEVSL) and income elasticity of non-fatal health effects literatures. If there are additional relevant empirical studies that should also be included in the summary, please provide citations.*

The SAB finds that Robinson and Hammitt (2015) and the EPA document *Technical Memorandum: Income Elasticity* provide reasonable summaries of the income elasticity literature. The SAB does, however, recommend that the EPA consider including the study by Murphy and Topel (2006) and the meta-analysis by Mrozek and Taylor (2002) in the summary. ~~Even i~~If these studies are not included in the EPA analysis, the agency should provide justification for not including the studies because they provide information that should be relevant. The SAB generally finds ~~is~~ that very little ~~(not enough)~~ research has been conducted in this important area. The EPA should support more research to provide methodological guidance and empirical estimates of the income elasticity of VSL. One area to explore further, in the absence of explicit studies, is the possibility of using estimates of the income elasticity for other related goods and services to infer estimates of the income elasticity of VSL (e.g., Chetty 2006 Hall and Jones 2007). Examples of related goods and services to consider for this purpose could include consumer products that can be used to reduce health risks such as bottled water and suntan lotion and various forms of health insurance. ***<<Chair's note: it would be helpful to provide some examples and citations to clarify what types of goods and services>>*** While this may not be straightforward, the ability to use such estimates would greatly increase the empirical basis upon which to ground the income elasticity of VSL.

Key Recommendations

- Robinson and Hammitt (2015) and the EPA document *Technical Memorandum: Income Elasticity* provide reasonable summaries of the income elasticity literature. However, the EPA should consider including the study by Murphy and Topel (2006) and the meta-analysis by Mrozek and Taylor (2002) in the summary.

- Very little research has been conducted on the income elasticity of the value of statistical life. The EPA should support more research to provide methodological guidance and empirical estimates in this important area.

- In the absence of explicit studies, the EPA should support research that may enable~~consider the possibility of~~ the use~~ing~~ of estimates of the income elasticity for other related goods and services (such as consumer products that can be used to reduce health risks and various forms of health insurance) to infer estimates of the income elasticity of the value of statistical life.

3.5.2. Analysis of Very Low Income Elasticity Estimates

Charge Question 14. Several reported mean income elasticity estimates from stated preference studies are quite low, sometimes even zero. The "balanced" approach in the EPA Technical Memorandum does not include reported mean estimates of zero, but does include very low reported mean estimates (e.g., 0.1). Please comment on whether this an appropriate and scientifically sound choice. How should very low, non-zero, mean reported income elasticity results be addressed in the analysis?

The SAB finds that it is highly ~~unlikely~~implausible for the income elasticity of VSL to be zero or negative. However, it is not completely clear how such estimates should be addressed in the EPA's analysis. It can be argued that such estimates may be theoretically impossible and therefore should be dropped, but it is hard to find statistical justification for dropping them. One statistical justification for dropping them, however, is that the income in these studies was not well measured, which may bias the estimates toward zero. Imprecision in the quality of measurement will be partly reflected in the standard errors of the individual income elasticity of statistical life estimates. Perhaps some of these estimates will not pass the stricter validity tests that will be imposed as discussed in the response to Charge Question 1a. This may render these points moot.

Commented [K44]: Issue of identification of structural paprameters and criteria for dropping when we consider estimates an approximation to a theoretical concept there is no basis for dropping estimates.

The SAB recommends that the EPA adopt the following strategies:

1. Instead of calculating an unweighted mean of income elasticity of VSL estimates, use standard errors of individual income elasticity of VSL estimates to calculate a weighted mean.

2. Given the lack of congruence on the low/zero estimates, calculate the weighted mean of the income elasticity of VSL both with and without the low/zero estimates to assess their influence.

Key Recommendations

- Instead of calculating an unweighted mean of income elasticity of VSL estimates, the EPA should use standard errors of individual income elasticity of VSL estimates to calculate a weighted mean.

- Given the lack of congruence on the low/zero income elasticity of VSL estimates, the EPA should calculate the weighted mean of the income elasticity of VSL both with and without the low/zero estimates to assess their influence

3.5.3. Study Selection Criteria and Alternative Approaches for Estimating Central Income Elasticity of Value of Statistical Life

Charge Question 15. Please comment on whether the selection criteria applied by Robinson and Hammitt (2015) are clearly enumerated, appropriate, and scientifically sound and whether the additional inclusion of Viscusi, Huber, and Bell (2014) in the Technical Memorandum is appropriate based on results reported in the study's on-line appendix (attached).

Charge Question 16. Given the relatively limited number of studies upon which to draw for estimating the income elasticity of VSL, the EPA Technical Memorandum describes two alternatives for arriving at a central income elasticity of VSL estimate and range for use in environmental policy analysis. Of these alternatives which is the most appropriate and

scientifically sound? Please provide the rationale for your choice. Would it be appropriate to consider using the alternative as a sensitivity or uncertainty characterization?

Charge questions 15 and 16 pertain to the same general topic, how to best arrive at an estimate of the income elasticity of the VSL. These charge questions are therefore discussed together.

EPA's Selection Criteria and Alternatives for Estimating Income Elasticity of VSL

The SAB finds that neither of the two alternatives put forward in Robinson and Hammitt (2015) and described in EPA's technical memorandum represent an adequate basis for providing an estimate(s) of the income elasticity of VSL for policy purposes. With regard to the first option, using the central estimates and range from a meta-analysis, Robinson and Hammitt (2015) do an admirable job summarizing the available literature. Their analysis, however, drives home the point that there is not an adequate informational basis for deriving a consensus estimate of the income elasticity of VSL. The inclusion or non-inclusion of the Viscusi, Huber and Bell (2014) does not alter this conclusion. Robinson and Hammitt's (2015) inclusion of studies that are publically available, but not in the peer reviewed literature clashes with the EPA study selection criteria used for determining a central estimate for the VSL, but is best seen as an indication of the lack of an adequate information basis for estimating a central value for the income elasticity of VSL. The second option that Robinson and Hammitt (2015) put forward is to use estimates from the Viscusi (2015) meta-analysis of hedonic pricing results that rely on the CFOI data. This meta-analysis is recent and was performed competently but the set of studies used is somewhat narrow. The preferred estimates from this study are substantially larger than those found in other recent meta-analyses that draw on broader set of studies, including those by Lindhjem, et al. (2011) and Doucouliagos, Stanley and Viscusi (2014) which suggest much lower central values for the income elasticity of VSL.

Nature of the Problem Faced in Estimating Income Elasticity of VSL

It is useful to understand several aspects of the nature of the problem faced in arriving at an estimate of the income elasticity of VSL for policy purposes.

1. To estimate the income elasticity of VSL, variation in income is needed. However, there has been relatively little change in median income over the last two decades particularly for groups represented in the samples used for hedonic wage studies. Changes in per capita income have been more pronounced, but much of the change has been in the two tails of the income distribution. This calls into question what the appropriate income variable is if a causal relationship is needed.

2. Some studies estimate the income elasticity of VSL from a cross section of individuals while others estimate the income elasticity of VSL from time series data. It is well known that estimates based on cross sectional data measure what would be expected to happen to an individual's VSL if that individual swapped income with someone else in the current income distribution. In contrast, income elasticity of VSL estimates based on a time series measure provide an estimate of how VSL statistics would shift if the entire income distribution rises or falls. The EPA's use of income elasticity of VSL estimates to adjust VSL estimates over time generally calls for a time series-based measure.

3. The hedonic wage approach does not, by design, provide an estimate of the income elasticity of VSL.[6]

4. While stated preference studies are carefully designed to produce reliable VSL estimates, this is not the case for income elasticity of VSL estimates. The single income question asked in the typical stated preference survey is most often taken from standard government surveys and its initial use is to help make a determination as to whether the data collected are adequately representative of the population of interest with respect to income. This is done by comparing the distribution in income to that of U.S. Census Bureau statistics. This type of income question is known to be fraught with measurement error due to substantial respondent heterogeneity with respect to what constitutes income and to suffer from having a high rate of missing values.[7] It has long been known that in order to adequately measure income, a very large set of questions about specific types of income and monetary transfers is required.[8] Furthermore, from a theoretical perspective, income is not the correct variable that should help determine the risk-wage tradeoff but rather the correct variable is medium term discretionary wealth. The best that can be hoped for is that a simple regression of this variable on income, as typically measured in surveys, has independent and identically distributed normal error terms. In this case, the presence of classical measurement error is known to bias the estimate of the income elasticity of VSL downward, a result that has considerable support in the broader literature on income elasticities.

Methodologies for Estimating Income Elasticity of VSL

Smith and Evans (2010) identify four methodologies to estimate the income elasticity of VSL: (1) stated preference studies; (2) meta-analyses of hedonic wage studies; (3) cross-country comparisons of VSL estimates; and (4) comparisons of VSL estimates at different points in time for a single country. Robinson and Hammitt (2015) concentrate on the first two. The two main problems with the stated preference estimates of the income elasticity of VSL were noted previously: they are cross-sectional estimates rather than time series estimates and they suffer from substantial measurement error problems with respect to income. A meta-analysis of hedonic wage studies might serve as a basis on which to estimate the income elasticity of VSL. However, to make this work one needs a large number of studies across time periods with both income variation and a relatively constant mix of estimation techniques used to estimate the VSL in those different time periods. Unfortunately, there are not a large number of available studies and the desire of journals to publish papers using new methodologies means that particular methodologies for estimating the VSL are always confounded with time/income variation.[9] Using cross-country comparisons of VSL estimates is an attempt to increase the range of income levels

[6] The use of quantile regression, e.g., Kniesner et al. (2010) and Evans and Schaur (2010), to estimate a hedonic wage equation can potentially provide a cross-sectional estimate of the income elasticity of VSL at different points in the wage distribution if there is wage-related heterogeneity in the wage-risk tradeoffs being made by individuals in the sample.

[7] A common example here is that some retired people view drawing money from a retirement savings account like an IRA to be income while others don't.

[8] For the two exemplars of purpose built that do this, see the Survey of Consumer Finances sponsored by the U.S. Federal Reserve Board and the U.S. Census Bureau's Survey of Income and Program Participation.

[9] It would also be desirable to have a number of distinct data sources among the studies used the meta-analysis that were evenly distributed over time periods with different income. Unfortunately, the available studies often share some common data sources but are idiosyncratic enough with respect how key variables are constructed that these differences too are confounded with the specific time period when the study was conducted.

observed and hence to be able to statistically estimate the income elasticity of VSL with reasonable precision. There are several difficulties with this approach. The preferences of people in other countries may be systematically different from people living in the United States. Indeed, this is the rationale advanced by the EPA for not relying on VSL estimates in other countries. A variant of the cross sectional data problem is seen when considering the situation where the different VSL estimates used in estimating the income elasticity of VSL come from different countries in the same year.

The fourth approach of comparing VSL estimates at different points in time from a single country provides a coherent way to obtain an income elasticity of VSL estimate for policy purposes. This approach might be useful for providing insights into the plausibility of income elasticity of zero versus one, whether or not the analysis provides a compelling point estimate of income elasticity. An example of this approach is found in Costa and Kahn (2004) who look at the evolution of the VSL from 1940 to 1980. Their work is not relevant to the EPA's current need because their analysis stops in 1980 and the Census of Fatal Occupational Injuries ~~Cost of Funds Index~~ (CFOI~~COFI~~) risk data being used in current hedonic wage studies does not exist for the time period Costa and Kahn examine. It would be possible, however, to take one of the currently preferred VSL model specifications that can be estimated by combining the U.S. Census Bureau's Annual Social and Economic Supplement to the Current Population Survey (CPS) with ~~COFI~~CFOI data.[10] By holding the methodology and data sources used to estimate the VSL constant, it should be possible to use the income variation over the last two decades to obtain a defensible income elasticity of VSL estimate.[11] Each annual cross section of the CPS, can be used to produce a VSL estimate. To each of these VSL estimates, the desired measure of income for that year can be attached. Calculation of the income elasticity of VSL is then a straightforward econometric exercise. The sensitivity of the income elasticity of VSL estimate to the different model specifications for estimating the VSL can be examined and the resulting income elasticity of VSL estimates averaged if there is not a clear reason for favoring one model specification over another. The sensitivity of the income elasticity of VSL estimate to the particular definition of income can also be examined. For example, income elasticity of VSL estimates could be estimated using median per capita income and gross domestic product (GDP) per capita. The income elasticity of VSL estimate(s) to be used in assessing regulations could be updated at regular intervals simply by adding VSL estimates based on more recent years of the CPS, with earlier time period perhaps given less weight in determining the income elasticity of VSL estimate.

Key Recommendations

- Neither of the two alternatives put forward in Robinson and Hammitt (2015) and described in EPA's technical memorandum represent an adequate basis for providing an estimate(s) of the income elasticity of VSL for policy purposes. Therefore the SAB recommends that the EPA consider an alternative approach.

Commented [K45]: Need work on what can be identified

[10] Some of hedonic wage regressions use the ~~COFI~~CFOI rates averaged over multiple years. Doing this is similar to including a lagged regressor the sense of reducing the effective number of observations in the regression model by the length of the lag period.

[11] Much of the effort would be in the form of preparing the CPS and ~~COFI~~CFOI data for the first cross-sectional hedonic wage regression. Because subsequent cross-sections would use the same variable definitions and industry-occupation fatality rates, the data preparation and program effort involved should be substantially reduced.

- Comparing VSL estimates at different points in time from a single country provides a coherent way to obtain an income elasticity of VSL estimate for policy purposes. The SAB recommends selecting one of the currently preferred VSL model specifications that can be estimated by combining the U.S. Census Bureau's Annual Social and Economic Supplement to the Current Population Survey (CPS) with ~~COFI~~CFOI data and using the income variation over the last two decades to obtain a defensible income elasticity of VSL estimate.

- The SAB recommends examining the sensitivity of the income elasticity of VSL estimate to different model specifications and averaging the resulting income elasticity of VSL estimates if there is not a clear reason for favoring one model specification over another.

3.5.4. Income Elasticity of the Value of Non-fatal Health Effects

> *Charge Question 17. As described in Robinson and Hammitt (2015), there are limited data on income elasticity of non-fatal health effects. As a result the Technical Memorandum recommends using the income elasticity of VSL to estimate income elasticity for the value of these non-fatal health risks. Please comment on whether this represents an appropriate and scientifically sound approach given the available data.*

The SAB recognizes that there are limited data available on income elasticity of non-fatal health effects but does not ~~fully~~ support using the income elasticity of VSL to estimate income elasticity for the value of these non-fatal health risks as an interim solution. The SAB finds that it is conceptually incorrect to apply income elasticity for one good to some other good, even though the two goods are related in some way. Moreover, both the magnitude of the valuation estimates and the income elasticities seem likely to be influenced by changes in national health insurance policies.

The SAB recommends that the EPA explore the income elasticity of expenditures on private health care products as a better proxy for the income elasticity of non-fatal health risks. <<***Chair's note: can we provide citations?***>>

Key Recommendations

- The SAB does not ~~fully~~ support using the income elasticity of VSL to estimate income elasticity for the value of fatal health risks because it is conceptually incorrect to apply income elasticity for one good to some other good. The SAB recommends that the EPA consider using ~~explore~~ the income elasticity of expenditures on private health care products as a better proxy for the income elasticity of non-fatal health risks.

REFERENCES

Alberini A., M. Cropper, A. Krupnick, and N.B. Simon. 2004. Does the value of a statistical life vary with age and health status? Evidence from the U.S. and Canada. *Journal of Environmental Economics and Management* 48(1):769-792.

Aldy, J.E., and W.K. Viscusi. 2008. Adjusting the value of a statistical life for age and cohort effects. *The Review of Economics and Statistics* 90(3):573-581.

Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman. 1993. Report of the NOAA Panel on Contingent Valuation. *Federal Register* 58:4601-4614.

Ashenfelter, O., and M. Greenstone. 2004. Using mandated speed limits to measure the value of a statistical life. *Journal of Political Economy* 112(2, Part2):S226-S267

Bellavance, F., G. Dionne, and M. Lebeau. 2009. The value of a statistical life: A meta-analysis with a mixed effects regression model. *Journal of Health Economics* 28:444-464.

Belsley, D. 1991. *Conditional Diagnostics: Collinearity and Weak Data in Regression*. Wiley Series in Probability. John Wiley, New York, NY.

Belsley, D.A., E. Kuh, and R.E. Welsch. 1980. *Regression Diagnostics. Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, New York, NY.

Bishop, R.C., and K.J. Boyle. 2016. "Reliability and Validity in Nonmarket Valuation." Chapter 14 in *A Primer on Nonmarket Valuation*, P.A. Champ, K.J. Boyle and T.C. Brown (eds). Springer.

Boyle, K.J. 2003. "Contingent Valuation in Practice." Chapter 14 in *A Primer on Nonmarket Valuation*, P.A. Champ, K.J. Boyle and T.C. Brown (eds). Springer.

Bowland, B.J., J.C. Beghin. 2001. Robust estimates of the value of a statistical life for developing countries. *Journal of Policy Modelling* 23:385-396.

Cameron, T.A., and J.R. DeShazo. 2009. *Demand for Health Risk Reductions*. Department of Economics, University of Oregon Working Paper.

Cameron, T.A., and J.R. DeShazo. 2013. Demand for health risk reductions. *Journal of Environmental Economics and Management* 65:87-109.

Cameron, T.A., J.R. DeShazo, and E.H. Johnson. 2010. The effect of children on adult demands for health-risk reductions. *Journal of Health Economics* 29:364-376.

Cameron, T.A., J.R. DeShazo, and P. Stiffler. 2013. Demand for health risk reductions: A cross-national comparison between the U.S. and Canada. *Journal of Risk and Uncertainty* 41:245-273.

Carmines, E.G., and R.A. Zeller. 1979. *Reliability and validity assessment.* Vol 17. Sage publications.

Carson, R., N.E. Flores, and W.M. Hanemann. 1998. Sequencing and valuing public goods. *Journal of Environmental Economics and Management* 36 (3):314-323.

Carson, R.C., T. Groves, and J.A. List. 2014. Consequentiality: a theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists* 1(1):171-207.

Chestnut, L.G., R.D. Rowe, and W.S Breffle. 2012. Economic valuation of mortality-risk reduction: stated preference estimates from the United States and Canada. *Contemporary Economic Policy* 30(3):399-416.

Chetty, R. 2006. A new method of estimating risk aversion. *American Economic Review* 96(5):1821-1834.

Corso, P.S., J.K. Hammitt, and J.D. Graham. 2001. Valuing mortality-risk reduction: using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty* 23(2):165-184.

Costa, D.L., and M.E. Kahn. 2004. Changes in the value of life, 1940--1980, *Journal of Risk and Uncertainty* 29(2):159-180.

Davis, L. 2004. The effect of health risk on housing values: evidence from a cancer cluster. *American Economic Review* 94(5):1693-1704

Davison, A.C. and D.V. Hinkley. 1997. *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge, UK.

De Blaeij, A., R.J.G.M. Florax, P. Rietveld, and E. Verhoef. 2003. The value of statistical life in road safety: A meta-analysis. *Accident Analysis and Prevention* 35:973-986.

Deleire, T, S. Khan, and C. Timmins. 2013. Roy model sorting and the value of a statistical life. *International Economic Review* 54(1):279-306.

Doucouliagos, H., T.D. Stanley, and W.K. Viscusi. 2014. Publication selection and the income elasticity of the value of a statistical life. *Journal of Health Economics* 33(2014):67-75.

Evans, M.F., and G. Schaur. 2010. A quantile Estimation approach to identify income and age variation in the value of a statistical life. *Journal of Environmental Economics and Management* 59(3): 260-270.

Gentry, E.P., and W. K Viscusi, 2016. The fatality and morbidity components of the value of statistical life. *Journal of Health Economics* 46(March 2016):90-99.

Hall, R.E. and C.I. Jones. 2007. The Value of life and the rise in health spending. *The Quarterly Journal of Economics* 122(1):39-72. doi: 10.1162/qjec.122.1.39.

Hammitt, J.K., and J.D. Graham. 1999. Willingness to pay for health protection: inadequate sensitivity to probability. *Journal of Risk and Uncertainty* 8:33-62.

Hammitt, J.K., and K. Haninger. 2010. Valuing fatal risks to children and adults: effects of disease, latency, and risk aversion. *Journal of Risk and Uncertainty* 40:57-83.

Hauber, A.B., A.O. Fairchild, and F.R. Johnson. 2013. Quantifying benefit-risk preferences for medical interventions: an overview of a growing empirical literature. Applied Health Economics and Health Policy 11(4):319-29.

Hedges, L.V., and I. Olkin. 1985. *Statistical Method for Meta-Analysis.* Academic Press.

Hsiang, S.M., M. Burke, and E. Miguel. 2013. Quantifying the influence of climate on human conflict. *Science* 341:1235367.

Hunter, J.E., and F.L. Schmidt. 2004. *Methods of Meta-Analysis*. Sage Publications.

Kniesner, T.J., W.K. Viscusi, and J.P. Ziliak. 2010. Policy relevant heterogeneity in the value of statistical life: New evidence from panel data quantile regressions. *Journal of Risk and Uncertainty* 40(1): 15–31.

Kochi, I., B. Hubbell, and R. Kramer. 2006. An empirical Bayes approach to combining and comparing estimates of the value of statistical life for environmental policy analysis. *Environmental and Resource Economics* 34:385-406.

Lindhjem, H., S. Navrud, N.A. Braathen, and V. Biausque. 2011. Valuing mortality risk reductions from environmental, transport, and health policies: A global meta-analysis of stated preference studies. *Risk Analysis* 31: 1381-1407.

Manyika, J., M. Chui, J. Bughin, R. Dobbs, P. Bisson, and A. Marrs. 2013. *Disruptive technologies: Advances that will transform life, business, and the global economy*. McKinsey Global Institute.

Miller, T.R. 2000. Variations between countries in values of statistical life. *Journal of Transport Economics and Policy* 34(2):169-188.

Moher, D., A. Liberati, J. Tetzlaff, and D.G. Altman. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine* 151(4):264-269.

Mrozek, J.R., and L.O. Taylor. 2002. What determines the value of life? A meta-analysis. *Journal of Policy analysis and Management* 21(2): 253-270.

Murphy, K., and R. Topel. 2006. The value of health and longevity. *Journal of Political Economy* 114: 871-904.

Nelson, J.P., and P.E. Kennedy. 2009. The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environmental and Resource Economics* 42(3):345-377.

OECD (Organisation for Economic Co-operation and Development) 2112. *Mortality Risk Valuation in Environment, Health and Transport Policies*. OECD Publishing Paris, France. [Available at: http://www.oecd-ilibrary.org/environment/mortality-risk-valuation-in-environment-health-and-transport-policies_9789264130807-en]

Ren, S., H. Lai, W. Tong, M. Aminzadeh, X. Hou, and S. Lai. 2010. Nonparametric bootstrapping for hierarchical data. *Journal of Applied Statistics* 37(9):1487-1498.

Robinson, L.A. 2008. *Valuing Mortality Risk Reductions in Homeland Security Regulatory Analyses*. Final Report to U.S. Customs and Border Protection, U.S. Department of Homeland Security under subcontract to Industrial Economics, Inc.

Robinson, L.A., and J.K. Hammitt. 2015. *The effect of income on the value of mortality and morbidity risk reductions*. Review draft prepared for U.S. EPA. Contract EP-D-14-032 with Industrial Economics, Inc.

Robinson, L.A., and J.K. Hammitt. 2015. *The effect of income on the value of mortality and morbidity risk reductions*. Review draft prepared for U.S. EPA. Contract EP-D-14-032 with Industrial Economics, Inc.

Smith, V.K., G.L. Van Houtven, and S.K. Pattanayak. 2002. Benefit transfer via preference calibration: prudential algebra for Policy. *Land Economics* 78 (1):132–152.

Smith, V.K., S.K. Pattanayak, and G. Van Houtven. 2003. VSL reconsidered: what do labor supply estimates tell us about risk preferences? *Economics Letters*, 80:147–153.

Smith, V. K., K.S.K. Pattanayak, and G.L. Van Houtven. 2006. Structural benefit transfer: An example using VSL estimates. *Ecological Economics* 60(2):361-371.

Sunstein, C.R. 1997. Bad deaths. *Journal of Risk and Uncertainty*. 14:259-282.

The White House 2016. *Open Government Directive*. https://www.whitehouse.gov/open/documents/open-government-directive

Tufts University Medical Center. 2016. *Cost Effectiveness Analysis Registry*. https://research.tufts-nemc.org/cear4/Default.aspx. [Accessed, July 23, 2016]

U.S. EPA. 2010. *Guidelines for Preparing Economic Analyses*. National Center for Environmental Economics, Office of Policy, U.S. Environmental Protection Agency (updated May 2014).

U.S. EPA Science Advisory Board. 2011. *Review of Valuing Mortality Risk Reductions for Environmental Policy: A White Paper (December 10, 2010)*. EPA-SAB-11-011. U.S. Environmental Protection Agency Science Advisory Board, Washington, D.C. [Available at: http://yosemite.epa.gov/sab/sabproduct.nsf/298E1F50F844BC23852578DC0059A616/$File/EPA-SAB-11-011-unsigned.pdf]

U.S. Department of Health and Human Services. 2015. Patient Preference Information – Submission, Review in PMAs, HDE Applications, and De Novo Requests, and Inclusion in Device Labeling: Draft Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders. Draft Guidance. May 18, 2015. Food and Drug Administration, Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. [Available at: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM446680.pdf.]

U.S. General Services Administration. 2016. *How to Get Your Open Data on Data.gov.* http://www.digitalgov.gov/resources/how-to-get-your-open-data-on-data-gov/

U.S. Office of Management and Budget and U.S. Office of Science and Technology Policy. 2016. *Project Open Data.* https://project-open-data.cio.gov/

Van Houtven, G., F.R. Johnson, V. Kilambi, and A.B. Hauber. 2011. Eliciting benefit−risk preferences and probability-weighted utility using choice-format conjoint analysis. *Medical Decision Making* 31(3):469-480.

Viscusi, W.K. 2004. The value of life: Estimates with risks by occupation and industry. *Economic Inquiry* 42(1):29-48.

Viscusi, W.K. 2015. The role of publication selection bias in estimates of the value of a statistical life. *American journal of Health Economics* 1(1):27-52.

Viscusi, W.K., and J.E. Aldy. 2003. The value of a statistical life: A critical review of market estimates throughout the world. *The journal of Risk and Uncertainty* 27(1):5-76.

Viscusi, W.K., and J.E. Aldy. 2007. Labor market estimates of the senior discount for the value of statistical life. *Journal of Environmental Economics and Management* 53(3):377-392.

Viscusi, W.K., J. Huber, and J. Bell. 2014. Assessing whether there is a cancer premium for the value of a statistical life. *Health Economics* 23:384-396.

Viscusi, W.K., J. Huber, and J. Bell. 2014. Assessing whether there is a cancer premium for the value of a statistical life. *Health Economics* 23:384-396. [On-line appendix available at: http://onlinelibrary.wiley.com/doi/10.1002/hec.2919/suppinfo]

Viscusi, W. K., and E.P. Gentry. 2015. The value of a statistical life for transportation regulations: A test of the benefits transfer methodology. *Journal of Risk and Uncertainty* 51(1):53-77

APPENDIX A: THE EPA'S CHARGE QUESTIONS

Charge Questions for SAB-EEAC Review of an EPA White Paper: "Valuing Mortality Risk for Environmental Policy: a Meta-analytic Approach" and Technical Memorandum: "Income Elasticity of VSL"

February 2016

White Paper: Meta-analysis dataset

The White Paper assembles a database of stated preference and hedonic wage estimates of the value of statistical life (VSL) and, where possible, their standard errors. Criteria for inclusion in the database are based on recommendations from the SAB-EEAC (U.S. EPA Science Advisory Board 2011) (see section 4.4, page 13-20). EPA requests comments on whether the selection criteria previously recommended by the SAB-EEAC were appropriately interpreted and applied both for selecting studies to include in the meta-analysis and for selecting estimates within studies. **In answering questions 1(a) – 1(c), in addition to responding to the specific questions, please comment, in general, on whether the selection criteria previously recommended by the SAB-EEAC have been appropriately interpreted and applied in the White Paper.**

1a. Evidence of validity for stated preference studies: The SAB noted in its earlier advisory report (U.S. EPA Science Advisory Board 2011) that each selected stated preference study "should provide evidence that it yields valid estimates" (page 16). The SAB did not, however, specify how validity should be assessed. In applying this criteria, EPA included studies and estimates that passed a weak scope test or provided other evidence of validity (e.g., a positive coefficient on the risk variable as in the appendix for Viscusi, Huber and Bell 2014) as explained in Appendix B of the White Paper. Please comment on whether the methods EPA used in the White Paper to assess the validity of studies and estimates are appropriate and scientifically sound.

1b. Construct of the risk variable in hedonic wage studies: The SAB noted in its earlier advisory that the EPA should "Eliminate any study that relies on risk measures constructed at the industry level only (not by occupation within an industry)" (U.S. EPA Science Advisory Board 2011, page 18). It is not clear whether the SAB's parenthetical addition was meant as an example or as a directive. Only four studies constructed the risk variable by occupation and industry and met other selection criteria. In applying this criteria EPA included studies and estimates where the risk measure is differentiated by industry and at least one other characteristic (e.g., occupation, gender, age). Please comment on whether the hedonic wage studies included in the White Paper constructed the risk variable in a manner appropriate for use in the meta-analysis.

1c. Estimates for immediate risk reductions: To estimate the average value of the marginal willingness to pay for reduced risk of immediate death, the EPA selected estimates from the Stated Preference literature that are most closely comparable to the accidental deaths from the hedonic wage literature. The EPA made several judgement calls in determining the appropriate estimates to use from the stated preference literature. Specifically, Viscusi, Huber and Bell (2014) estimate reductions in risk of bladder cancer that will occur in 10 years. The authors discount the estimates to derive a comparable estimate for an immediate risk reduction. Alberini, et al. (2004) estimate a willingness to pay for an annual

reduction in risk over 10 years. We include estimates from both of these studies in the meta-analysis. Please comment on whether appropriate estimates from the stated preference literature were used in the White Paper to estimate the marginal willingness to pay for reduced risk of immediate death.

2. Please comment on whether relevant empirical studies in the stated preference and hedonic wage literatures are adequately captured in the White Paper. If additional studies should be included in the white Paper please provide citations.

3. Some estimates in the meta-analysis dataset in the White Paper are constructed by weighting subpopulation-specific estimates within a study in order to approximate an estimate for the general population. The specific weights used are described in Appendix B of the White Paper. Please comment on whether the population-weighting approach used in the White Paper is appropriate and scientifically sound.

4. In some cases EPA estimated standard errors in the White Paper using information within studies or provided by the study authors, as described in Appendix B. Please comment on whether the methods used in the White Paper to estimate standard errors when such information was not readily available are appropriate and scientifically sound.

White Paper: Analysis

Section 4 of the White Paper describes methods used to estimate representative VSL estimates from the meta-analysis dataset and presents results.

5. Please comment on whether the methodology used in the White Paper to analyze the data represents an appropriate and scientifically sound application of meta-analytic methods to derive generally applicable VSL estimates for environmental policy analysis.

6. The White Paper classifies estimates into independent samples, also called groups, as described in Section 4. Estimates from some hedonic wage studies that use the same or very similar worker samples are grouped together for the analysis. Similarly, some of the stated preference estimates using the same sample are grouped together. Please comment on whether this methodology represents an appropriate and scientifically sound approach for accounting for potential correlation of results that rely on the same underlying data.

7. Section 4.1 of the White Paper presents an expression that characterizes optimal weights that account for sampling and non-sampling errors, a framework that guides EPA's approach. Please comment on whether this is an appropriate and scientifically sound approach for addressing sampling and non-sampling errors.

8. The analysis in the White Paper adopts both non-parametric and parametric approaches (sections 4.1 and 4.2, respectively). Please comment on whether these approaches span a reasonable range of appropriate, scientifically sound, and defensible approaches to estimating a broadly applicable VSL for environmental policy and whether there are other methods that are more appropriate than those used in the White Paper.

White Paper: Results

9. The White Paper presents estimates using parametric and non-parametric models, pooled across stated preference and hedonic wage studies as well as balanced (i.e., equal weight to each study type), and weighted using different approaches. Of the range of estimates presented (see Section 4) the White Paper proposes the use of estimates from the following models:
    - Non-parametric model, balanced, mean of study mean
    - Parametric, balanced

Please comment on whether these proposed estimates represent reasonable and scientifically sound conclusions from the analyses in the White Paper and whether there is a different set (or sets) of results that are preferable based on the data and analysis in the White Paper.

10. The results section of the White Paper concludes with an influence analysis. Please comment on whether this analysis is a reasonable way to characterize the influence of individual studies on the estimated VSLs, whether the results of the influence analysis suggest any changes or modifications to the estimation approach, and whether it is important to include an influence analysis.

Establishing a Protocol for Future Revisions:

11. In the previous SAB advisory report (USEPA Science Advisory Board 2011), the SAB endorsed the idea of establishing a standardized protocol and regular schedule for future updates to the Agency's mortality risk valuation estimates. Please comment on relevant statistical criteria for the inclusion of additional eligible estimates and/or the exclusion of older estimates that could help inform the development of a standardized protocol for future updates and the timing or frequency of those updates.

12. In its 2011 report the SAB-EEAC recommended "…EPA work toward developing a set of estimates…for policy-relevant cases characterized by risk…" (U.S. EPA Science Advisory Board 2011, pp. 10). Among the studies that meet the selection criteria in the current White Paper, three stated preference studies provide values for reductions in risks of cancer (i.e., Hammitt and Haninger 2010, Chestnut, Rowe, and Breffle 2012, and Viscusi, Huber and Bell 2014). Only two of those studies (Hammitt and Haninger 2010 and Chestnut, Rowe, and Breffle 2012) allow for a within study comparison of values for cancer and non-cancer risk reductions. However, EPA could augment the literature by modifying the selection criteria to include studies from other countries or from the grey literature, and/or using other methods (e.g., risk-risk studies). Please comment on whether, and if so how, selection criteria for identifying studies for estimating a cancer differential should differ from those used in the current White Paper. Does the literature support a non-zero cancer differential?

Technical Memorandum: Income elasticity

13. The EPA document *Technical Memorandum: Income Elasticity* presents a summary of the recent income elasticity literature based on a review presented in Robinson and Hammitt (2015). Please comment on whether Robinson and Hammitt (2015) and the EPA Technical Memorandum provide an appropriate and scientifically sound summary of the income elasticity of VSL (IEVSL) and income elasticity of non-fatal health effects literatures. If there are additional relevant empirical studies that should also be included in the summary, please provide citations.

14. Several reported mean income elasticity estimates from stated preference studies are quite low, sometimes even zero. The "balanced" approach in the EPA Technical Memorandum does not include reported mean estimates of zero, but does include very low reported mean estimates (e.g., 0.1). Please comment on whether this an appropriate and scientifically sound choice. How should very low, non-zero, mean reported income elasticity results be addressed in the analysis?

15. Please comment on whether the selection criteria applied by Robinson and Hammitt (2015) are clearly enumerated, appropriate, and scientifically sound and whether the additional inclusion of Viscusi, Huber, and Bell (2014) in the Technical Memorandum is appropriate based on results reported in the study's on-line appendix (attached).

16. Given the relatively limited number of studies upon which to draw for estimating the income elasticity of VSL, the EPA Technical Memorandum describes two alternatives for arriving at a central IEVSL estimate and range for use in environmental policy analysis. Of these alternatives which is the most appropriate and scientifically sound? Please provide the rationale for your choice. Would it be appropriate to consider using the alternative as a sensitivity or uncertainty characterization?

17. As described in Robinson and Hammitt (2015), there are limited data on income elasticity of non-fatal health effects. As a result the Technical Memorandum recommends using the IEVSL to estimate income elasticity for the value of these non-fatal health risks. Please comment on whether this represents an appropriate and scientifically sound approach given the available data.

References

Alberini, A., M. Cropper, A. Krupnick, and N.B. Simon. 2004. Does the value of a statistical life vary with age and health status? Evidence from the US and Canada. *Journal of Environmental Economics and Managemen*t 48(1):769-792.

Chestnut, L.G., R.D. Rowe, and W.S. Breffle. 2012. Economic valuation of mortality-risk reduction: stated preference estimates from the United States and Canada. *Contemporary Economic Policy* 30(3):399-416.

Hammitt, J.K., and K. Haninger. 2010. Valuing fatal risks to children and adults: effects of disease, latency, and risk aversion. *Journal of Risk and Uncertainty* 40:57-83.

Robinson, L.A., and J.K. Hammitt. 2015. *The effect of income on the value of mortality and morbidity risk reductions*. Review draft prepared for U.S. EPA. Contract EP-D-14-032 with Industrial Economics, Inc.

U.S. EPA Science Advisory Board. 2011. *Review of Valuing Mortality Risk Reductions for Environmental Policy: A White Paper (December 10, 2010)*. Office of the Administrator, Science Advisory Board. EPA-SAB-11-011. July 29. [Available at: http://yosemite.epa.gov/sab/sabproduct.nsf/298E1F50F844BC23852578DC0059A616/$File/EPA-SAB-11-011-unsigned.pdf]

Viscusi, W.K., J. Huber, and J. Bell. 2014. Assessing whether there is a cancer premium for the value of a statistical life. *Health Economics* 23:384-396. [On-line appendix available at: http://onlinelibrary.wiley.com/doi/10.1002/hec.2919/suppinfo]

APPENDIX B BIBLIOGRAPHY ON WILLINGNESS TO PAY IN HEALTH AND HEALTH CARE [PARTIAL]

Bala, M.V., LL. Wood, G.A. Zarkin, E.C. Norton, A. Gafni, and B. O'Brien. 1998. Valuing outcomes in health care: a comparison of willingness to pay and quality-adjusted life-years. *Journal of Clinical Epidemiology* 51(8):667-76. doi: S0895-4356(98)00036-5 [pii].

Blumenschein K., and M. Johannesson. 1996. Economic evaluation in health care: A brief history and future directions. *Pharmacoeconomics* 10(2):114-12.

Blumenschein K., and M. Johannesson. 1996. Incorporating quality of life changes into economic evaluations of health care: an overview. *Health Policy* 36:155-166.

Blumenschein, K., and M. Johannesson,. 1998. Relationship Between Quality of Life Instruments, Health State Utilities, and Willingness to Pay in Patients with Asthma. Annals of Allergy, Asthma and Immunology 80(2):189-94.

Blumenschein K., M. Johannesson, K. Yokoyama, and P. Freeman. 2001. Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. *Journal of Health Economics* 20: 441-457

Bosworth, R., T.A. Cameron, and J.R. DeShazo. 2006. Preferences for preventative public health policies with jointly estimated rates of time preference. School of Public Health and International Affairs, North Carolina State University, Raleigh, NC.

Cameron, T.A., and J.R. Deshazo. 2005. Valuing health-risk reductions: Sick-years, lost life-years, and latency. CCPR-053-05, *On-Line Working Paper Series, California Center for Population Research*. [Available at: http://papers.ccpr.ucla.edu/papers/PWP-CCPR-2005-053/PWP-CCPR-2005-053.pdf]

Chestnut L.G., L.R. Keller, W.E. Lambert, and R.D. Rowe. 1996. Measuring heart patients' willingness to pay for changes in angina symptoms. *Medical Decision Making* 16:65-77.

Corso, P.S., J.K. Hammitt, and J.D. Graham. 2001. Valuing mortality-risk reduction: Using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty* 23(2):165-184.

Diener, A., B. O'Brien, and A. Gafni. 1998. Health care contingent valuation studies: A review and classification of the literature. *Health Economics* 7:313-326.

Fisman, D.N., M.A. Mittleman, G.S. Sorock, and A.D. Harris. 2002. Willingness to pay to avoid sharps-related injuries: A study in injured health care workers. *American Journal of Infection Control* 30(5): 283-287.

Gan, T.J., F. Sloan, G.L. Dear, H.E. El-Moalem, and D.A. Lubarsky. 2001. How much are patients willing to pay to avoid postoperative nausea and vomiting? *Anesthesia and Analgesia* 92(2): 393-400.

Gyrd-Hansen, D., P.A. Halvorsen, and I.S. Kristiansen. 2008. Willingness-to-pay for a statistical life in the times of a pandemic. *Health Economics* 17:55-66.

Henson, S. 1996. Consumer willingness to pay for reductions in the risk of food poisoning in the UK. *Journal of Agricultural Economics* 47(3):403-420.

Hirth, R.A., M.E. Chernew, E. Miller, A.M. Fendrick and W.G. Weissert. 2000. Willingness to pay for a quality-adjusted life year: In search of a standard. *Medical Decision Making* 20:332-342.

Jacobs, R.J., R.J. Moleski, and A.S. Meyerhoff. 2002. Valuation of symptomatic hepatitis A in adults. *Pharamacoeconomics* 20(11):739-747.

Johnson, F.R. 2005. Einstein on willingness to pay per QALY: Is there a better way? *Medical Decision Making* 25:607-608.

Johnson, F.R., W.H. Desvouges, M.C. Ruby, D. Stieb, and P. De Civita. 1998. .Eliciting stated health preferences: An application to willingness to pay for longevity. *Medical Decision Making* 18 suppl: S57-S67.

Johnson, F.R., M.R. Banzhaf, and W.H. Desvousges. 2000. Willingness to pay for improved respiratory and cardiovascular health: A multiple-format stated-preference approach. *Health Economics* 9(4):295-317.

Johnson, F.R., R. Manjunath, C.A. Mansfield, L.J. Clayton, T.J. Hoerger and P. Zhang. 2006. High-risk individuals' willingness to pay for diabetes risk-reduction programs. *Diabetes Care* 29:1351-1356.

Jones-Lee, M.W. 1993. Personal willingness to pay for prevention: evaluating the consequences of accidents as a basis for preventive measures. *Addiction* 88:913-921.

Kartman, B., F. Andersson, and M. Johannesson. 1996. Willingness to pay for reductions in angina pectoris attacks. *Medical Decision Making* 16(3):248-253.

Kartman, B., N-O Stalhammar, and M. Johannesson. 1996. Valuation of health changes with the contingent valuation method: A test of scope and question order effects. *Health Economics* 5:531-541.

Keith, P.L., J. Haddon, and S. Birch. 2000. A cost-benefit analysis using a WTP questionnaire of intranasal budesonide for seasonal allergic rhinitis. *Annals of Allergy and Asthma Immunology* 84:5562.

Kenkel, D. 2006. WTP- and QALY-based approaches to valuing health for policy: Common ground and disputed territory. *Environmental and Resource Economics* 34:419-437.

Kleinman, L., E. McIntosh, M. Ryan, J. Schmier, J. Crawley, G.R. Locke, and G. De Lissovoy. 2002. Willingness to pay for complete symptom relief of gastroesophageal reflux disease. *Archives of Internal Medicine* 162:1361-1366.

Krabbe, P.F., M.L. Essink-Bot, and G. Bonsel. 1997. The comparability and reliability of five health-state valuation methods. *Social Science and Medicine* 45(11):1641-52.

Krupnick, A.J., and M. Cropper. 1992. The effect of information on health risk valuations. *Journal of Risk and Uncertainty* 5(1):29-48.

Krupnick, A., A. Alberini, M. Cropper, N. Simon, B. O'Brien, R. Goeree, and M. 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* 24:161-186.

Liu Jin-Tan, J.K. Hammitt, and J.D. Wang. 2000. Mother's WTP for her own and her child's health: a contingent valuation study in Taiwan. *Health Economics* 9:319-26.

Luchini, S., C. Protiere C, and J.P. Moatti. 2002. Eliciting several willingness to pay in a single contingent valuation survey: application to health care. *Health Economics* 12(1):51-64.

Lundberg, L., M. Johannesson, M. Silverdahl, and M. Lindberg. 1999. Quality of life, health-state utilities and WTP in patients with psoriasis and atopic eczema. *British Journal of Dermatology* 141(6):1067-75.

Maddala, T., K.A. Phillips, and F.R. Johnson. 2003. An experiment on simplifying conjoint analysis designs for measuring preferences. *Health Economic*s 12(12):1035–1047.

Magnusson, M. 1996. *Theory and Methods of Economic Evaluation of Health Care*. Dordrecht: Kluwer Academic Publishers.

Marshall, D.A., F.R. Johnson, K.A. Phillips, J.K. Marshall, L. Thabane, and N.A. Kulin. 2007. Measuring patient preferences for colorectal cancer screening using a choice-format survey. *Value in Health* 10:415-430.

Morrison, G.C., and M. Gyldmark. 1992. Appraising the use of contingent valuation. *Health Economics* 1:233-243.

Navrud, S. 2001. Valuing health impacts from air pollution in Europe. *Environmental and Resource Economics* 20:305-329.

Norinder, A, K. Hjalte, and U. Persson. 2001. Scope and scale insensitivities in a contingent valuation study of risk reductions. *Health Policy* 57:141-153

O'Brien, B.J., and J.L. Viramontes. 1994. Willingness to pay: A valid and reliable measure of health state preferences? *Medical Decision Making* 14(3):289-297.

O'Conor, R.M., and G.C. Blomquist. 1997. Measurement of consumer-patient preferences using a hybrid contingent valuation method. *Journal of Health Economics* 16:667-683.

Olsen, J. A. 1997. Aiding priority setting in health care: Is there a role for the contingent valuation method? *Health Economics* 6:603-612.

Olsen, J.A., and R.D. Smith. 2001. Theory versus practice: A review of 'Willingness to Pay' in health and health care. *Health Economics* 10:39-52.

Perreira, K.M., and F. Sloan. 2002. Living healthy and living long: Valuing the nonpecuniary loss from disability and death. *The Journal of Risk and Uncertainty* 24(1):5-29.

Phillips, K.A., T. Maddala, and F.R. Johnson. 2002. Measuring preferences for health care interventions using conjoint analysis: An application to HIV testing. *Health Services Research* 37:1681-1705.

Read, D., and N.L. Read, 2001. An age-embedding effect: time sensitivity and time insensitivity when pricing health benefits. *Acta Psychologica* 108:117-136.

Russell, S., J. Fox-Rushby, and D. Arhin. 1995. Willingness and ability to pay for health care: a selection of methods and issues. *Health Policy Planning* 10(1): 94-101.

Ryan, M., D.A. Scott, and C. Donaldson. 2004. Valuing health care using willingness to pay: A comparison of the payment card and dichotomous choice methods. *Journal of Health Economics* 23: 237-258.

Ryan, M., D.A. Scott, C. Reeves, A. Bate, E.R. van Teijlingen, E.N. Russel, M. Napper, and C.M. Robb. 2001. Eliciting public preferences for health care: a systematic review of techniques. *Health Technology Assessment* 5(5):1-186.

Sloan, F.A., W.K. Viscusi,, H.W. Chesson, C.J. Conover, and K. Whetten-Goldstein. 1998. Alternative approaches to valuing intangible health losses: The evidence for multiple sclerosis. *Journal of Health Economics* 17:475-497.

Slothuus, U., and R.G. Brooks. 2000. WTP in arthritis: a Danish contribution. *Rheumatology* (Oxford) 39(7):791-9.

Slothuus, U., M.L. Larsen, and P. Junker. 2000. WTP for arthritis symptom alleviation. *International Journal of Technology Assessment in Health Care* 16 (1): 60-72.

Slothuus, U., M. Larsen, and P. Junker. 2001. The contingent ranking method—a feasible and valid method when eliciting preferences for health care? *Social Science and Medicine* 54:1601-1609.

Smith, R. 1999. *Exploring the Relationship Between Time Trade-off and Willingness-to-Pay: An Empirical Investigation*. Centre for Health Program Evaluation. Working Paper 88.

Smith, R.D. 2001. The relative sensitivity of willingness-to-pay and time trade-off to changes in health status: An empirical investigation. *Health Economics* 10:487-497.

Stavem, K. 1999. WTP: A feasible method for assessing treatment benefits in epilepsy? *Seizure* 8:14-19.

Stavem, K. 2002. Association of willingness to pay with severity of chronic obstructive pulmonary disease, health status and other preference measures. *International Journal of Tuberculosis and Lung Disease* 6(6):542-549.

Stewart, J.M., E. O'Shea, C. Donaldson, and P. Shackley. 2002. Do ordering effects matter in willingness to pay studies of health care? *Journal of Health Economics* 21(4):585-599.

Telser, H., and P. Zweifel. 2002. Measuring willingness-to-pay for risk reduction: an application of conjoint analysis. *Health Economics* 11(2):129-139.

Thompson, M.S. 1986. Willingness to pay and accept risks to cure chronic disease. *American Journal of Public Health* 76(4):392-396.

Tolley, G., D. Kenkel, and R. Fabian. 1994. *Valuing Health for Policy: An Economic Approach*. The University of Chicago Press, Chicago.

Viscusi, W.K., W.A. Magat, and J. Huber. 1991. Pricing health risks: Survey assessments of risk-risk and risk-dollar tradeoffs. *Journal of Environmental Economics and Management* 21(1):32-51.

Watson, V., and M. Ryan. 2007. Exploring preference anomalies in double bounded contingent valuation. *Journal of Health Economics* 26:463-482.

Zethraeus, N. 1998. Willingness to pay for hormone replacement therapy. *Health Economics* 7: 31-38.

Zillich, A.J., K. Blumenschein, M. Johannesson, and P. Freeman. 2002. Assessment of the relationship between measures of disease severity, quality of life, and willingness to pay in Asthma. *Pharmacoeconomics* 20(4): 257-265.

APPENDIX C: BIBLIOGRAPHY ON BENEFIT-RISK AND RISK-RISK TRADEOFF PREFERENCES IN HEALTH AND HEALTH CARE [PARTIAL]

Arden, N.K., A.B. Hauber, A.F. Mohamed, F.R. Johnson, P.M. Peloso, D.J. Watson, P. Mavros, A. Gammaitoni, S.S. Sen, and S.D. Taylor. 2012. How do physicians weigh benefits and risks associated with treatments in patients with osteoarthritis in the United Kingdom? *Journal of Rheumatology* 39(5):1056-63.

Barker, J.H., A. Furr, M. Cunningham, F. Grossi, D. Vasilic, B. Storey. O. Wiggins, R. Majzoub, M. Vossen, P. Brouha, C. Maldonado, C.C. Reynolds, C. Francois, G. Perez-Abadia, J.M. Frank, M. Kon, and J.C. Banis Jr. 2005. Investigation of risk acceptance in facial transplantation. *Plastic and Reconstructive Surgery* 118(3):663–70.

Bewtra, M., and F.R. Johnson. 2013. Assessing patient preferences for treatment options and process of care in inflammatory bowel disease: A critical review of quantitative data. *Patient* 6 (4) 241-55.

Bewtra, M., V. Kilambi, A.O. Fairchild, C.A. Siegel, J.D. Lewis, and F.R. Johnson. 2014. Patient Preferences for Surgical versus Medical Therapy for Ulcerative Colitis. *Inflammatory Bowel Disease* 20(1):103-114.

Bremnes, R.M., K. Andersen, and E.A. Wist. 1995. Cancer patients, doctors and nurses vary in their willingness to undertake cancer chemotherapy. *European Journal of Cancer* 131A(12):1955–9.

Bridges, J.F., A.F. Mohamed, H.W. Finnern, A. Woehl, and A.B. Hauber. 2012. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. *Lung Cancer* 77(1):224–31.

Cook, R.D. and S. Weisberg. 1982. *Residuals and Influence in Regression*. Chapman and Hall, New York, NY.

Cross, J, J. Yang, F.R. Johnson, J. Quiroz, J. Dunn, M. Raspa, and D.B. Bailey. 2016. Caregiver preferences for the treatment of males with fragile X syndrome. *Journal of Developmental and Behavioral Pediatrics* 37:71-79.

de Bekker-Grob, E.W., M.L. Essink-Bot, W.J. Meerding, H.A. Pols, B.W. Koes, and E.W. Steyerberg. 2008. Patients' preferences for osteoporosis drug treatment: a discrete choice experiment. *Osteoporosis International* 19(7):1029–37.

Devereaux, P.J., D.R. Anderson, M.J. Gardner, W. Putnam, G.J. Flower-dew, B.F. Brownell, S. Nagpal, and J.L. Cox. 2001. Differences between perspectives of physicians and patients on anticoagulation in patients with atrial fibrillation: observational study. *BMJ* 323(7323):1218–22.

Eberth, B., V. Watson, M. Ryan, J. Hughes, and G. Barnet. 2009. Does one size fit all? Investigating heterogeneity in men's preferences for benign prostatic hyperplasia treatment using mixed logit analysis. *Medical Decision Making* 29(6):707–15.

Flynn, T.N., J.J. Louviere, T.J. Peters, and J. Coast. 2007. Best-worst scaling: What it can do for health care research and how to do it. *Journal of Health Economics* 26 (1):171-89.

Fraenkel, L, S. Bodardus, and D.R. Wittnik. 2001. Understanding patient preferences for the treatment of lupus nephritis with adaptive conjoint analysis. *Medical Care* 39(11):1203–16.

Fraenkel, L, B. Gulanski, and D. Wittink.2007. Patient willingness to take teriparatide. P*atient Education and Counseling* 65(2):237–44.

Gafni, A., and G.W. Torrance. 1984. Risk Attitude and Time Preference in Health. *Management Science* 30(4):440-451. doi: doi:10.1287/mnsc.30.4.440.

Gallego, G., J.F. Bridges, T. Flynn, B.M. Blauvelt, and L.W. Niessen. 2012. Using best-worst scaling in horizon scanning for hepatocellular carcinoma technologies. *International Journal of Technology Assessment in Health Care* 28 (3):339-46.

Gonzalez, J.M., F.R. Johnson, M.C. Runken, and C.M. Poulos. 2013. Evaluating migraineurs' preferences for migraine treatment outcomes using a choice experiment. *Headache* 53:1635–1650.

Gonzalez, J.M., A.B. Hauber, and F.R. Johnson. 2015. Estimating conditional certainty equivalents using choice-experiment data. *Journal of Choice Modelling* 15:14-25.

Hauber, A.B., F.R. Johnson, H. Fillit, A.F. Mohamed, C. Leibman, H.M. Arrighi M. Grundman, and R.J. Townsend. 2009. Older Americans' risk-benefit preferences for modifying the course of Alzheimer disease. *Alzheimer Disease and Associated Disorders* 23(1):23–32.

Hauber, A.B., A.F. Mohamed, M.E. Watson, F.R. Johnson, and J.E. Hernandez. 2009. Benefits, risk, and uncertainty: preferences of antiretroviral-naive African Americans for HIV treatments. *AIDS Patient Care and STDs* 23(1):29–34.

Hauber, A.B., F.R. Johnson, K.M. Grotzinger, and S. Ozdemir. 2010. Patients' benefit–risk preferences for chronic idiopathic thrombocytopenic purpura therapies. *Annals of Pharmacotherapy* 44(3):479–88.

Hauber, A.B, N.K. Arden, A.F. Mohamed, F.R. Johnson, P.M. Peloso, D.J. Watson, P. Mavros, A. Gammaitoni, S.S. Sen, and A.D. Taylor. 2013. A discrete-choice experiment of United Kingdom patients' willingness to risk adverse events for improved function and pain control in osteoarthritis. *Osteoarthritis Cartilage* 21(2):289-97.

Hauber, A.B., A.O. Fairchild, and F.R. Johnson. 2013. Quantifying benefit-risk preferences for medical interventions: an overview of a growing empirical literature. *Applied Health Economics and Health Policy* 11(4):319-29.

Hauber, A.B., A.F. Mohamed, F.R. Johnson, M. Cook, H.M. Arrighi, J. Zhang, and M. Grundman. 2014. Understanding the relative importance of preserving functional abilities in Alzheimer's disease in the United States and Germany. *Quality of Life Research* 23(6):1813-1821.

Ho, M.P., J.M. Gonzalez, H.P. Lerner, C.Y. Neuland, J.M. Whang, M. McMurry-Heath, A.B. Hauber, and T. Irony. 2015. Incorporating patient-preference evidence into regulatory decision making. *Surgical Endoscopy* (10):2984-93. doi: 10.1007/s00464-014-4044-2. Epub 2015 Jan 1.

Holden, W.L. 2003. Benefit–risk analysis: a brief review and proposed quantitative approaches. *Drug Safety* 26(12):853–62.

Hollin, I. L., H. L. Peay, and J. F. Bridges. 2015. Caregiver preferences for emerging Duchenne muscular dystrophy treatments: a comparison of best-worst scaling and conjoint analysis. *Patient* 8 (1):19-27. doi: 10.1007/s40271-014-0104-x.

Johnson, F.R., S. Ozdemir, B. Hauber, and T.L. Kauf. 2007. Women's willingness to accept perceived risks for vasomotor symptom relief. *Journal of Women's Health.* 16(7):1028–40.

Johnson, F.R., S. Ozdemir, C. Mansfield, S. Hass, D.W. Miller, C.A. Siegel, and B.E. Sands. 2007. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. *Gastroenterology* 133(3):769–79.

Johnson, F.R., S. Ozdemir, C. Mansfield, S. Hass, C.A. Siegel, and B.E. Sands. 2008. Are adult patients more tolerant of treatment risks than parents of juvenile patients? *Risk Analysis* 29(1):121–36.

Johnson, F.R., G. Van Houtven, S. Ozdemir, S. Hass, J. White, G. Francis, D.W. Miller, and J.T. Phillips. 2009. Multiple sclerosis patients' benefit–risk preferences: serious adverse event risks versus treatment efficacy. *Journal of Neurology* 256(4):554–62.

Johnson, F.R., A.B. Hauber, S. Ozdemir, and L. Lynd. 2010. Quantifying women's stated benefit–risk trade-off preferences for IBS treatment outcomes. *Value in Health.* 13(4):418–23.

Johnson, F.R., B. Hauber, S. Ozdemir, C.A. Siegel, S. Hass, and B.E. Sands. 2010. Are gastroenterologists less tolerant of treatment risks than patients? Benefit–risk preferences in Crohn's disease management. *Journal of Managed Care Pharmacy* 16(8):616–28.

Johnson, F.R., A.B. Hauber, and J. Zhang. 2013. Quantifying Patient Preferences to Inform Benefit-Risk Evaluations. In Benefit-Risk Analysis in *Pharmaceutical Research and Development*, A Sashegyi, J. Felli and B. Noel (eds.) New York: Chapman & Hall.

Kauf, T.L., J.C. Yang. A.B. Kimball, M. Sundaram, Y. Bao, M. Okun, P. Mulani, A.B. Hauber, amd F.R. Johnson. 2015. Psoriasis patients' willingness to accept side-effect risks for improved treatment efficacy *Journal of Dermatological Treatment* 26(6):507-13.

Kopec, J.A., C.G. Richardson, H. Llewellyn-Thomas, A. Klinkhoff, A. Carswell, and A. Chalmers. 2007. Probabilistic threshold technique showed that patients' preferences for specific trade-offs between pain relief and each side effect of treatment in osteoarthritis varied. *Journal of Clinical Epidemiology* 60(9):929–38.

Levitan, B., M. Markowitz, A.F. Mohamed, F.R. Johnson, L. Alphs, L.L. Citrome, and J.F. Bridges. 2015. Patients' preferences related to benefits, risks, and formulations of schizophrenia treatment. *Psychiatric Services* 66(7):719-26 DOI: 10.1176/appi.ps.201400188 [Epub]

Lewis, S.M., F.N. Cullinane, A.J. Bishop, L.S. Chitty, T.M. Marteau, and J.L. Halliday. 2006. A comparison of Australian and UK obstetricians' and midwives' preferences for screening tests for down syndrome. *Prenatal Diagnosis* 26(1):60–6.

Llewellyn-Thomas, H.A., J.I. Williams, L. Levy, and C.D. Naylor. 1996. Using a trade-off technique to assess patients' treatment preferences for benign prostatic hyperplasia. Medical Decision Making 16(3): 262–82.

Llewellyn-Thomas, H.A., J.M. Paterson, J.A. Carter, A. Basinski, M.G. Myers, G.D. Hardacre, E.V. Dunn, R.B. D'Agostino, and C.C. Naylor. 2002. Primary prevention drug therapy: can it meet patients' requirements for reduced risk? *Medical Decision Making* 22:326–39.

Lynd, L.D., M. Naiafzadeh, L. Colley, M.F. Byrne, A.R. Willan, M.J. Sculpher, F.R. Johnson, and A.B. Haube. 2010. Using the incremental net benefit framework for quantitative benefit–risk analysis in regulatory decision making—a case study of alosetron in irritable bowel syndrome. *Value in Health*. 13(4):411–7.

Majzoub, R.K., M. Cunningham, F. Grossi, C. Maldonado, J.C. Banis, and J.H. Barker. 2006. Investigation of risk acceptance in hand transplantation. J*ournal of Hand Surgery (American Volume)* 31(2):295–302.

Manjunath, R., J.C. Yang, and A.B. Ettinger. 2012. Patients' preferences for treatment outcomes of add-on antiepileptic drugs: a conjoint analysis. *Epilepsy and Behavior* 24(4):474-9.

Markowitz, M., B. Levitan, A.F. Mohamed, F.R. Johnson, J.F.P. Bridges, L. Alphs, and L. Citrome. 2014. Psychiatrists' judgments around antipsychotic benefit and risk outcomes and formulation in schizophrenia treatments. *Psychiatric Services* 65(9):1133-9.

McTaggart-Cowan, H.M., P. Shi, J.M. Fitzgerald, A.H. Anis, J.A. Kopec, T.R. Bai, J.A. Soon, and L.D. Lynd. 2008. An evaluation of patients' willingness to trade symptom-free days for asthma-related treatment risks: a discrete choice experiment. *Journal of Asthma* 45(8):630–8.

Mohamed, A.F., A.B. Hauber, and M.P. Neary. 2011. Patient benefit–risk preferences for targeted agents in the treatment of renal cell carcinoma. *Pharmacoeconomics* 29(11):977–88.

Mohamed, A.F., F.R. Johnson, A.B. Hauber, B. Lescrauwaet, and M. Saylan. 2012. Do patients and physicians have similar preferences for chronic hepatitis B treatment outcomes in Turkey? *Flora: The Journal of Infectious Diseases and Clinical Microbiology* 17(1):29-38.

Mohamed, A.F., Z. Zhang, F.R. Johnson, I. Duprat Lomon, E. Malvolti, C. J. Ostgren, and K.G. Parhofer. 2013. The avoidance of weight gain is important for oral type 2 diabetes treatments in Sweden and Germany: Patient preferences. *Diabetes and Metabolism* 39:397-403.

Mohamed, A.F., F.R. Johnson, M-M Balp, and F. Calado. 2016. Preferences and stated adherence for antibiotic treatment of cystic fibrosis Pseudomonas infections. *The Patient - Patient-Centered Outcomes Research*. 9(1):59-67.

O'Brien, B.J., J. Elswood, and A. Calin. 1990. Willingness to accept risk in the treatment of rheumatic disease. *Journal of Epidemiology and Community Health* 44(3):249–52.

Palda, V.A., H.A. Llewellyn-Thomas, R.G. Mackenzie, K.I. Pritchard, and C.D. Naylor. 1997. Breast cancer patients' attitudes about rationing postlumpectomy radiation therapy: applicability of trade-off methods to policy-making. *Journal of Clinical Oncology* 15(10):3192–200.

Pullar, T., V. Wright, and M. Feely. 1990. What do patients and rheumatologists regard as an 'acceptable' risk in the treatment of rheumatic disease? *British Journal of Rheumatology* 29(3):215–8.

Ratcliffe, J., and M. Buxton. 1999. Patients' preferences regarding the process and outcomes of life-saving technology. An application of conjoint analysis to liver transplantation. *International Journal of Technology Assessment in Health Care* 15(2):340–51.

Ratcliffe, J., M. Buxton, T. McGarry, R. Sheldon, and J. Chancellor. 2004. Patients' preferences for characteristics associated with treatments for osteoarthritis. *Rheumatology (Oxford)* 43(3): 337–45.

Reynolds, C.C., S.A. Martinez, A. Furr, M. Cunningham, J.M. Bumpous, E.J. Lentsch, J.C. Banis, D. Vasilic, B. Storey, O. Wiggins, C. Maldonado, G. Perez-Abadia, and J.H. Barker. 2006. Risk acceptance in laryngeal transplantation. *Laryngoscope* 116(10):1770–5.

Richardson, C.G., A. Chalmers, H.A. Llewellyn-Thomas, A. Klinkhoff, A. Carswell, and J.A. Kopec. 2007. Pain relief in osteoarthritis: patients' willingness to risk medication-induced gastrointestinal, cardiovascular, and cerebrovascular complications. *Journal of Rheumatology* 34(7):1569–75.

Salkeld, G., M. Solomon, L. Short, M. Ryan, and J.E. Ward. 2003. Evidence-based consumer choice: a case study in colorectal cancer screening. *Australian and New Zealand Journal of Public Health* 27(4):449–55.

Sassi, F., and M. McKee. 2008. Do clinicians always maximize patient outcomes? A conjoint analysis of preferences for carotid artery testing. *Journal of Health Services Research and Policy* 13(2):61–6.

Simes, R.J., and A.S. Coates. 2001. Patient preferences for adjuvant chemotherapy of early breast cancer: how much benefit is needed? *Journal of the National Cancer Institute Monographs* 30:146–52.

Van Houtven, G., F.R. Johnson, V. Kilambi, and H.B. Hauber. 2011. Eliciting benefit–risk preferences and probability-weighted utility using choice-format conjoint analysis. *Medical Decision Making* 31(3): 469–80.

Wong, M.K., A.F. Mohamed, A.B. Hauber, J.C. Yang, Z. Liu, J. Rogerio, and C.A. Garay. 2012. Patients rank toxicity against progression free survival in second-line treatment of advanced renal cell carcinoma. *Journal of Medical Economics* 15(6):1139–48.

Yang, J.C., F.R. Johnson, V. Kilambi, and A.F. Mohamed. 2015. Sample size and utility-difference precision in discrete-choice experiments: A meta-simulation approach. *Journal of Choice Modelling* 16:50-57. DOI: 10.1016/j.jocm.2015.09.001. Epub 11/3/2015.

Zhang, J., F.R. Johnson, A.F. Mohamed, and A.B. Hauber. 2015. Too many attributes: A test of the validity of combining discrete-choice and best-worst scaling data. *Journal of Choice Modelling* 15:1-13.