Synthesizing Research for Benefit Transfer: Valuing Mortality Risk Reductions

Lisa A. Robinson (Harvard University),* James K. Hammitt (Harvard University), Kevin Haninger (U.S. Department of Health and Human Services)

Working Paper prepared for:

Methods for Research Synthesis:
A Cross-Disciplinary Workshop

Harvard Center for Risk Analysis

October 3, 2013

www.hcra.harvard.edu

***Corresponding author: robinson@hsph.harvard.edu**

Disclaimer: The findings and conclusions of this paper are those of the authors and do not imply endorsement by any component of Harvard University or other sponsors of this workshop, nor do they necessarily represent the views of the U.S. Department of Health and Human Services. Comments should be directed to the authors.

Acknowledgements: Dr. Hammitt's and Ms. Robinson's work was funded in part by National Science Foundation grant 1227327and U.S. Department of Agriculture agreement 58-5000-3-0035.

Synthesizing Research for Benefit Transfer: Valuing Mortality Risk Reductions

Lisa A. Robinson (Harvard University), James K. Hammitt (Harvard University), Kevin Haninger (U.S. Department of Health and Human Services)

Abstract: When conducting benefit-cost analysis, analysts often find that the available benefit valuation research does not address the same risks and populations as addressed by the policy. In such cases, they typically use the benefit transfer framework to evaluate the existing research and select estimates to apply in the analysis. Although this framework is a type of research synthesis, it differs from the use of synthesis methods in many other settings. Rather than combining estimates from studies that address the phenomenon of concern, benefit transfer involves extrapolating from the studied context to another context. We explore related issues, focusing on how benefit transfer techniques are used to estimate the value of small changes in mortality risks in U.S. regulatory analyses, conventionally reported as the value per statistical life (VSL). We find that Federal agencies currently rely on point estimates from selected studies, in part because further development of meta-analysis and other synthesis methods is needed.

This need for methodological improvements makes it difficult to assess the extent to which the choice of synthesis method affects the benefit transfer results. The available meta-analyses result in central VSL estimates ranging from about $2 million to $9 million, but the selection criteria and analytic techniques they apply have been criticized by recent expert panels. The qualitative literature reviews that support the values used by government agencies, as well as individual “best practice” studies, result in values towards the high end of this range. While other synthesis methods, including structural models and expert elicitation, show promise, they have not yet progressed past the pilot phase.

We also find that which synthesis methods are most suitable in this context depends on the match between the available studies and the policy context, and on the extent to which refining the VSL estimate is likely to change the overall results of the benefit-cost analysis. In some cases, it may be appropriate to rely on estimates from an individual study or a group of studies accompanied by assessment of uncertainty; in others, more sophisticated analysis may be desired to better tailor the estimate to a particular policy context. Qualitative review or meta-analysis are most useful when the available research addresses populations and risks similar to

the policy context, or provides data that can be used to tailor the results to that context. Structural modeling and expert elicitation may be needed when theory and expert judgment are necessary to address gaps or inconsistencies in the available research. However, investment in more sophisticated analysis may not be worthwhile if better tailoring the VSL estimate is not likely to change the determination of whether a policy is cost-beneficial or of which option is likely to yield the largest net benefits, or if uncertainties in other parameter estimates have a more significant influence on the results. In such cases, it may be more important to invest in reducing uncertainty in other analytic components.

Keywords: benefit-cost analysis, value per statistical life, willingness to pay, health risk valuation

1.0 INTRODUCTION

Policy analysts are often confronted with difficult questions about whether and how to value benefits, such as reduced mortality risks that cannot be fully valued based on market prices. While well-established methods exist for nonmarket valuation, the available studies often address risks or populations that are somewhat dissimilar to those addressed by the policy under consideration. As a result, analysts must often choose among various options when conducting benefit-cost analysis, each of which has some undesirable consequences. The options include conducting new primary research that directly addresses the outcome of concern; excluding the outcome from the quantitative comparison of costs and benefits; or using various synthesis methods to apply the available research, describing the implications of its limitations.

The first option, conducting new primary research, is often infeasible. New valuation research requires substantial time and funding; typically policy decisions must be made quickly and the resources available to support analysis are limited. While these problems could be mitigated by implementing a research program that anticipates future needs, such programs are infrequently funded for valuation research and are not helpful for those analysts asked to address near-term policy decisions.

While feasible, the second option – excluding the outcome from the benefit-cost analysis – can make it more difficult to articulate and debate the rationale for the policy decision. Such exclusion may be particularly problematic if the outcome is expected to account for a substantial proportion of overall benefits or represents a major policy goal. Without some information on potential magnitude, decisionmakers may weight the outcome by an amount that is far greater or far less than the range of values suggested by the available evidence. If not reported in the analysis, such weighting may not be recognized and explicitly discussed as part of the decisionmaking process. In some cases, analysts will have no choice but to exclude the outcome from the quantitative analysis, if no information is available on its potential value. In many cases, however, some information is available and the question becomes how to best evaluate and apply it.

This brings us to the third option: applying available research that addresses similar outcomes. This option is generally referred to as “benefit transfer” and is frequently employed in benefit-cost analysis. It involves selecting an estimate or range of estimates, or developing a

valuation function, from the research literature and applying it in an analysis that addresses a somewhat different context. The basic steps include defining the outcome to be valued, searching the literature for relevant studies, evaluating the quality of the studies and assessing their applicability to the policy outcome, synthesizing the estimates based on qualitative review or quantitative methods such as meta-analysis, and assessing associated uncertainties.

The dividing line between direct application of primary research and benefit transfer is fuzzy. The distinction relates more to the extent of the differences between the study and policy scenario. Given the complexity of real-world policy settings, it is hard to imagine a valuation study that addresses a scenario that is identical in all respects to the policy context. In addition, any research involves some uncertainties whose direction and magnitude may be difficult to quantify. Such uncertainties may result from the methods, data, assumptions, and models used as well as from the need to simplify the context for valuation purposes. A well-conducted benefit transfer that addresses a somewhat dissimilar context, but relies on high quality primary research, may be preferable to a poorly-conducted primary research study. The benefit transfer framework requires analysts to explicitly address the differences between the study and policy scenario as well as the quality of the research, and to assess the implications of related uncertainties for decisionmaking.

Thus research synthesis is a fundamental component of the benefit transfer process. Evaluating the validity of the transfer is a difficult task, however, because the results generally cannot be compared to high quality research that directly addresses the outcome of concern. Comparison studies of other outcomes may provide useful insights, but may not be directly relevant due to differences in the approach and the valuation context. Well-informed professional judgment plays an important role, as does full disclosure of the approach, assumptions, results, and uncertainties.

In this paper, we explore these issues, focusing on how mortality risk reductions are valued in analyses of major U.S. environmental, health, and safety regulations. We address mortality risks because they often account for a significant fraction of the total quantifiable benefits of these policies and because the underlying research base is relatively well-established. The manner in which these estimates are synthesized has important policy implications and demonstrates the challenges analysts face even when substantial research is available. We begin

by briefly summarizing the benefit transfer framework then discuss issues related to application of synthesis methods in more detail.

2.0 BENEFIT TRANSFER AND RESEARCH SYNTHESIS

Methods for research synthesis, such as qualitative systematic review and quantitative meta-analysis, are typically used to combine the results of various studies to determine what is known about a particular quantity, when multiple studies exist that address the same phenomenon. In such cases, the goal is to better understand the phenomenon by integrating the findings from studies that offer different perspectives and have varying advantages and limitations. In the case of benefit transfer, we are also attempting to improve our understanding of a phenomenon – the value of a particular policy outcome – but lack studies that directly address that outcome. Instead, we look for insights from studies of similar outcomes. This section introduces the role of valuation in benefit-cost analysis then discusses best practice guidance for benefit transfer and issues related to determining its validity.

2.1 Valuation and Benefit-Cost Analysis

In benefit-cost analysis, the goal is to estimate how those affected value the policy outcome in monetary terms, so that benefits and costs are measured in the same units. Ultimately, the question is whether the benefits are likely to exceed the costs, and, if multiple policy options are under consideration, which option is likely to provide the largest net benefits. Benefit values, and the benefit-cost analysis they support, are typically a few among many inputs into policy decisions; not necessarily the determining factor.

As in other contexts, uncertainty and its implications may be addressed through qualitative discussion, sensitivity analysis, and/or probabilistic modeling. Breakeven or threshold analysis is also often useful. For example, if a policy option appears to result in net costs, analysts can estimate the increase in benefits needed for the policy to yield net benefits. Analysts and decisionmakers can inspect the results to determine whether the increase is within a range that appears plausible given what is known about benefit values. In some cases, large differences in benefit values may not affect the policy implications of the analysis; e.g., benefits may exceed costs by a substantial amount regardless of whether low- or high-end values are applied. In

others, small differences may matter greatly. Thus the implications of the uncertainties in a particular transfer will vary depending on the policy under consideration.

This varying importance means that different analyses will require different levels of investment in the valuation approach. Conducting more research may not necessarily change the conclusion that a particular option is (or is not) cost-beneficial or the ranking of the options. Ideally, analysts would use screening analysis, based on easily accessible data and simple assumptions, to provide preliminary information on the possible direction and magnitude of the effects and to inform decisions about future work.

Because time and resources are limited, the ideal analysis will not quantify all outcomes with equal precision. For some outcomes, the cost of further assessment will be greater than the likely benefits, given their importance for decisionmaking. In other words, devoting more resources to assessing the value of a particular outcome may not sufficiently improve the analysis to pass an informal benefit-cost or value of information test. Generally more attention should be paid to valuing outcomes that are likely to represent a significant fraction of the benefits (or costs), as well as outcomes that are of interest to decisionmakers and stakeholders regardless of their magnitude.

In this context, analysts face difficult questions about the level of resources to devote to fine-tuning the benefit transfer given that estimating the validity of the transfer is challenging. Thus determining the trade-off between the costs of increased investigation and the benefits of increased validity requires substantial judgment.

2.2 Benefit Transfer Guidance

We focus on how benefit transfer is applied in the analysis of major Federal environmental, health, and safety regulations. Such analysis is required under Executive Orders 12866 and 13543 (Clinton 1993, Obama 2011). The U.S. Office of Management and Budget (OMB), located within the Executive Office of the President, is responsible for reviewing analyses prepared by executive branch agencies and issued *Circular A-4: Regulatory Analysis* (OMB 2003) to provide government-wide guidance on their conduct. Of the agencies subject to these requirements, only the U.S. Environmental Protection Agency (EPA) has developed comprehensive guidance (EPA 2010a). This is not surprising given that EPA's air pollution

regulations dominate the set of regulations for which benefit-cost analysis is required (OMB 2013).

The OMB and EPA guidance address a wide range of topics, devoting only a few pages to benefit transfer. OMB (2003) focuses on criteria for evaluating the quality and applicability of the available research and discourages the use of benefit transfer given that such transfers involve "uncertainties and potential biases of unknown magnitude" (OMB 2003, p. 24).While acknowledging the concerns raised by OMB, EPA notes that "the reality is that benefit transfer is one of the most common approaches for completing a BCA [benefit-cost analysis] at EPA" (EPA 2010a, p. 7-45). Its attractiveness stems largely from the limited time and funding available; analysts rarely, if ever, have the resources needed to pursue new primary research.

The EPA (2010a) guidance describes the benefit transfer framework in more detail than the OMB (2003) guidance, including four steps summarized in Box 1.

Box 1: EPA Guidance on Benefit Transfer

Step 1: Describe the policy case.
Describe the characteristics and consequences of the policy, including the types of benefits it provides and the demographic and socioeconomic characteristics of the beneficiaries.

Step 2: Select study cases.
(a) Identify potentially relevant studies by searching the literature, including government reports, working papers, dissertations, unpublished research, and other "grey" literature as well as peer-reviewed journal articles.
(b) Develop selection criteria to evaluate the studies for quality and applicability. "Quality" should be assessed based on best practices for each valuation method. "Applicability" should be assessed by considering the similarity of the study case to the policy case in terms of (i) the commodity (i.e., outcome) being valued; (ii) the baseline and extent of change; and (iii) the characteristics of those affected.

Step 3: Transfer values.
Estimates may be transferred using point estimates from a single study or an average across several studies; a valuation function that adjusts for differences between the study case and the policy case; a meta-analysis that combines the results from several studies (which may range from simply ranking the results to more complex regression analysis); or, a structural model specified in accordance with theory and calibrated using estimates from the literature.

Step 4: Report the results.
In addition to reporting the results, analysts should describe key judgments and assumptions, including the criteria used to select study cases and the transfer approach. The uncertainty in the final benefit estimate should be quantified to the extent possible as well as discussed qualitatively.

Source: Authors' summary of EPA (2010a), pp. 7-45 to 7-49.

2.2 Benefit Transfer Validity

Determining the validity of a benefit transfer is challenging, because any valuation study involves some limitations and uncertainties. The principal means of testing validity is to compare a transfer to a primary research study that directly addresses the policy context. Several researchers have focused on whether the match between the transfer and the primary research results improves if the transfer is based on a function rather than a point estimate. Such functions involve developing an equation that relates the value to the selected characteristics of the population and context, based on a single study or (more frequently) meta-regressions that combine the results of several studies. Most such comparisons (e.g., Brouwer et al. 2000, Rosenberger and Loomis 2003, Johnson, Besedin, and Ranson 2006) find that a benefit transfer function performs better, but there is substantial variation in the results.

Rosenberger and Stanley (2006) discuss three potential sources of transfer errors. The first is generalization error, or the degree of similarity between study and policy context. Not surprisingly, they find that greater similarity reduces transfer error, as does the use of a function rather than a point estimate. The second is measurement error, or problems with the data, methods, and assumptions used in each study. They note that measurement error can be difficult to assess because many studies do not report all of the information needed. The third is publication selection bias, which results from the tendency of journals to accept articles that focus on innovation rather than apply best practices, as well as the tendency of both scholars and journals to not publish results that are statistically insignificant or inconsistent with expectations.

Another study (Kaul et al. 2013) uses a nonparametric approach to meta-analysis to explore 1,071 transfer errors reported in 31 comparison studies, and finds that the median absolute error is 39 percent, ranging from close to zero to over 200 percent. The authors note that the magnitude of these errors is similar to the magnitude of some types of errors found in primary research studies. Their findings on the sources of the error echo the Rosenberger and Stanley results noted above; they also find that transfers are less prone to error when used to predict the change in value associated with a change in quantity rather than quality.

These studies primarily consider transfers for environmental improvements (recreation, water quality, fishing, and so forth) rather than health effects. Only a few explore the validity of

transfers for mortality or morbidity risk reductions, and it is unclear whether such transfers follow similar patterns. For example, Ready et al. (2004) administer the same valuation survey across six European countries, focusing on morbidity risks. They then investigate how well the values for each country can be predicted based on the results for the five other areas. They find that some differences are correlated with differences in real income because benefit values generally increase with income, consistent with theoretical expectations (Hammitt and Robinson 2011). However, transferring values using different approaches – unit values with or without adjustment for income, or a valuation function that addresses demographic characteristics – perform equally well, with average transfer error of around 38 percent. Thus the magnitude of the error is similar to that found by Kaul et al., but the use of a transfer function appears to not significantly affect the results.

3.0 VALUING MORTALITY RISK REDUCTIONS

Benefit transfer is widely used to estimate the value of reductions in mortality risks associated with regulations and other policies. These risk reductions often represent a large proportion of the quantified benefits of environmental, health, and safety regulations, in many cases accounting for 90 percent or more of the benefits of the air pollution policies that dominate the U.S. regulatory agenda. Their value is typically expressed as the value per statistical life (VSL), defined as marginal rate of substitution between money and mortality risk in a short time period (e.g., a year).

Because these risk reductions are not bought and sold in the marketplace, they cannot be directly valued using market data; researchers instead rely on nonmarket valuation studies. These include stated preference studies that survey respondents about their willingness to pay (WTP) for the outcome of concern, including contingent valuation surveys and choice experiments. Alternatively, nonmarket outcomes can be valued using revealed preference methods. These methods rely on observed behaviors or prices and preferences for related market goods (such as the relationship between wages and job-related risks), and use statistical methods to separate the value of risk changes from the value of other attributes (for example, controlling for education when wages are used to estimate the value of job-related risks). Both methods have advantages and limitations. Perhaps most importantly, revealed preference studies have the advantage of

relying on actual behavior while stated preference methods rely on hypothetical payments, but the latter can be more easily tailored to particular policy contexts.

Conventionally, researchers first use these methods to develop estimates of individuals' WTP for a small change in their own risk within a defined time period. They then divide this WTP by the risk change to estimate the VSL (Hammitt 2000, Robinson and Hammitt 2013). For example, if an individual is willing to pay $900 for a 1 in 10,000 reduction in his or her risk of dying in the current year, his or her VSL is calculated as $900 WTP ÷ 1/10,000 risk change = $9.0 million. Because WTP is expected to change in proportion to the risk change as long as WTP is a small fraction of income, the resulting VSL can then be applied to small risk changes of varying magnitudes. The VSL is a convenient way to express individual WTP for small changes in one's own risk; it is not the value of saving an individual's life with certainty.

The application of the VSL in regulatory analysis faces several challenges related to the characteristics of the underlying research base as well as the context within which the estimates are applied. Below, we first introduce current practices. Then we highlight issues related to how research synthesis is applied in transferring VSL estimates for use in regulatory analysis.

3.1 Overview of Current Practices

OMB's 2003 guidelines allow regulatory agencies some discretion in selecting the VSL they apply in their analyses, indicating that the then-available research suggested that the VSL was generally between about $1 million and $10 million (2001 dollars). Of the major Federal agencies subject to the requirements for regulatory analysis, two – EPA and the Department of Transportation (DOT) – have implemented their own VSL guidance.[1]

Both agencies developed estimates in the early 1990s; however, EPA relied on a review by Viscusi (1992, 1993), while DOT relied on a review by Miller et al. (Urban Institute 1991), resulting in substantially different estimates. EPA's central VSL was $4.8 million in 1990 dollars, with a range from $0.6 million to $13.5 million, while DOT's was $2.5 million in 1992 dollars. While the agencies regulate different types of risks (i.e., largely illness-related deaths from exposure to environmental contaminants versus injury-related deaths from transportation accidents), the differing reviews and the values that result do not reflect attempts to match the

[1] A third agency, the Department of Homeland Security, uses a standard approach (based on Robinson 2008) but has not issued formal guidance.

primary research studies with the risks each regulates. Rather, due to limitations in the available research, both agencies relied on similar studies, most of which address job-related risks.[2,3] The difference is attributable to the studies selected and the extent to which the researchers adjusted the study estimates.

The estimates used by each agency have evolved over time. EPA's approach reflects substantial investment in primary research and consultation with expert panels, but has changed relatively little, while DOT has invested less in reviewing its approach but implemented more major changes. These differing processes reflect at least in part the extent to which the available research considers a context similar to the policy context. EPA's policies generally address illness-related risks to all age groups but particularly young and old, while the literature was long dominated by studies of injury-related risks that affect people of working age. DOT policies generally address injury-related risks and often affect a broad population.

EPA developed its approach in consultation with an independent Science Advisory Board expert panel (Schmalensee et al. 1993). It has since suggested several changes as part of its prospective analyses of the Clean Air Act, its development of agency-wide analytic guidance, and other activities, which were reviewed by a series of subsequent expert panels.[4] These efforts generally focus on two types of issues. The first relate to adjusting the base VSL for differences between the study scenarios and the regulatory scenarios. The second relate to updating the approach to incorporate more recent studies, including emerging meta-analyses. While the expert panels generally agreed that adjustments for scenario differences are desirable in principle, they have been reluctant to accept many of the suggested modifications. They frequently concluded that the available research literature was not sufficiently robust to support quantitative adjustments for various factors. Until recently, they also often suggested the available literature

[2] The original EPA review was completed in the early 1990s to support retrospective analysis of the Clean Air Act (EPA 1997) then incorporated into EPA's guidelines for regulatory analysis (EPA 2000a, EPA 2010a). Based Viscusi (1992, 1993), EPA identified 26 VSL estimates suitable for use in environmental policy analysis, of which 21 were from revealed preference studies that addressed job-related risks and five were from stated preference studies that relied on contingent valuation surveys. These surveys addressed job- or traffic-related risks.

[3] The report that underlies this DOT approach (Urban Institute 1991) indicates that the estimate is based on 47 VSL studies deemed to be "sound" in Miller (1990) plus two newer studies. Of the 49 studies, 30 consider the trade-off between wages and job-related risks, 11 address averting behavior or expenditures (e.g., for safer cars), and eight rely on surveys. The authors adjusted the estimates to reflect several shortcomings in the primary research studies.

[4] EPA's major proposals are discussed in EPA (2000b), Industrial Economics (2003), Dockins et al. (2004), Dockins et al. (2006), and EPA (2010). The expert panel reviews include Cropper et al. (1999), Stavins et al. (1999), Stavins et al. (2000), Cropper et al. (2001), Cameron et al. (2004), Cropper et al. (2007), Hammitt et al. (2010), and Kling et al. (2011). Dr. Hammitt was a member of several of these committees.

had not evolved sufficiently to warrant updating the base estimates.[5] However, in 2011, a Science Advisory Board committee (Kling et al. 2011) supported EPA's efforts to incorporate new research. EPA has not yet reported the results of implementing these recommendations nor updated its guidance.

Thus at present, EPA continues to rely on 26 point estimates identified from the Viscusi (1992, 1993) review to derive its base VSL (EPA 2000a, 2010a). EPA adjusts the base value for inflation and real income growth over time and discounts (using the same rate as applied elsewhere in the analysis) for any delay between changes in exposure and changes in mortality risk (i.e., latency or cessation lag).

In contrast, DOT has updated the set of studies used to support its estimates twice, in 2008 and 2013, resulting in larger VSL estimates.[6,7] It currently relies on nine wage-risk studies (DOT 2013), all of which were published in 2004 or later and hence not included in the group of studies that underlie EPA's estimates. Interestingly, the estimates used by the two agencies are now quite similar despite the differences in the underlying studies. In 2012 dollars, if adjusted only for inflation, EPA's central estimate is $8.4 million while DOT's is $8.9 million.[8] If also adjusted for real income growth (using the differing approaches applied by each agency), EPA's estimate becomes $9.2 million while DOT's is $9.1 million.[9]

[5] For a short period (roughly 2004 to 2007), EPA's Office of Air and Radiation (OAR) used the results of selected meta-analyses to estimate the VSL in their regulatory analyses, while other EPA offices continued to rely on the results from the Viscusi (1992, 1993) review. OAR applied a range of estimates anchored at $1 million (near the lower end of the range from Mrozek and Taylor 2002) and $10 million (near the upper end of the range from Viscusi and Aldy 2003), with a central estimate of $5.5 million (1999 dollars).

[6] In 2008, DOT revised its base estimate to $5.8 million, which was the simple mean of selected values from four meta-analyses (Miller 2000, Mrozek and Taylor 2002, Kochi et al. 2006, Viscusi and Aldy 2003), and one primary research study (Viscusi 2004). (The DOT guidance references earlier drafts of some of these analyses; we reference the published versions.)

[7] Ms. Robinson was a member of the expert advisory committee that supported development of the DOT (2013) guidance, and both Ms. Robinson and Dr. Hammitt subsequently served as peer reviewers of the draft guidance.

[8] Adjusted for inflation by the authors using the Consumer Price Index (CPI) inflation calculator at http://www.bls.gov/data/inflation_calculator.htm and base estimates reported in EPA (2010a) and DOT (2013). All estimates presented in the text are inflated to 2012 dollars, with no adjustment for real income growth, unless otherwise noted.

[9] EPA estimate updated by the authors, using adjustment factors provided by EPA. EPA applies a central income elasticity estimate of 0.4 while DOT applies an elasticity of 1.0; to estimate real income growth, EPA relies on data on gross domestic product (GDP) per capita while DOT relies on earnings data.

3.2 Study Selection and Evaluation

The EPA and DOT approaches summarized above, as well as the approaches used in other analyses, face several challenges related to identifying and evaluating relevant studies (i.e., step 2 in Box 1). Such steps form the foundation of any synthesis effort, since they are required no matter what approach is used for combining estimates. We address related issues below.

3.2.1 Reporting Bias

Reporting or publication bias is a fundamental problem for benefit transfer as well as for other forms of research synthesis. Peer-reviewed journals are less likely to accept best practice studies for publication, focusing instead on those that test new approaches. This both discourages the completion of the types of studies most needed for benefit transfer and means that those that are completed may be documented in the grey literature (e.g., in reports to funding agencies) rather than journal articles. In addition, both the researchers themselves and the journals may be reluctant to describe and publish findings that are statistically insignificant or do not conform to expectations.

Both EPA and DOT have, at times, included unpublished studies when developing their estimates, although most were subsequently published in the peer-reviewed literature. However, including unpublished literature, after appropriate review, does not necessarily alleviate the problems related to reporting bias. Regardless of whether the grey literature is included, it is impossible to include studies that were not conducted or research that was conducted but not documented.

The effect of reporting bias is difficult to assess or address, however. Doucouliagos, Stanley, and Giles (2012) explore its effects on meta-analyses of VSL estimates from wage-risk studies and conclude that the VSL may be significantly overstated. They note that this results at least in part because VSL researchers reject negative values as inconsistent with theory; i.e., a negative result implies that an individual would pay to increase, rather than decrease, his or her

risk of death.[10] However, negative values may instead result from measurement and sampling error.[11]

In benefit transfer, the bias that comes from excluding negative results may be counterbalanced by the tendency to exclude unusually large estimates. For example, DOT (2013) excludes studies that provide "implausibly large" estimates. Thus it is unclear whether the values used in benefit transfer are substantially smaller or larger than would be the case in the absence of reporting bias.

3.2.2 Research Quality

Regardless of what literature is included, peer review is not necessarily a substitute for review by the analysts conducting the transfer. The quality of peer review varies widely, and journal reviewers may focus on issues (such as the novelty of the approach) that differ from those that are important for benefit transfer (such as the representativeness of the sample). Benefit transfer requires the application of best practice standards that reflect evolving theory and empirical evidence. Thus familiarity with the subject matter and the methodology is needed, as is the exercise of informed professional judgment and clear disclosure of how this judgment was applied.

Best practices for revealed and stated preference studies are defined in general terms in the OMB (2003) and EPA (2010a) guidance; the standards for VSL studies are changing as the research evolves (see Cropper et al. 2011, EPA 2010b, and Kling et al. 2011). This progression reflects the development of new data sources and statistical techniques as well as new theoretical insights.

One example is the availability of more reliable data on job-related risks. In 1992, the Bureau of Labor Statistics (BLS) implemented the Census of Fatal Occupational Injuries (CFOI), which is a substantial improvement over the sources used in earlier VSL research (see Viscusi and Aldy 2003, Viscusi 2004 for more discussion). The CFOI involves extensive confirmation of the data and provides detailed information on each fatality. As a result, recent expert panels

[10] The authors note that it is also possible that researchers discard very large values that appear to be implausible outliers.

[11] Ashenfelter and Greenstone (2004) come to a similar conclusion, although their valuation methodology differs. They estimate the VSL in each U.S. state based on the relationship between the time saved by increased speed limits and fatality rates. If each state is treated as a separate study, rather than studied in combination with other states, dropping those states where the results lack statistical significance increases the estimated VSL.

(Kling et al. 2011, DOT 2013) recommended that, in developing their VSL estimates, EPA and DOT only include those studies that rely on risk data of equal or better quality than the CFOI. This eliminates all wage-risk studies published before 2003, including those that underlie the older guidance as well as many more recent meta-analyses.

Another example relates to concerns about individuals' ability to comprehend small risk changes in stated preference studies. Theory suggests that WTP for risk reduction should be nearly proportional to the risk reduction, as long as WTP is a small fraction of income. However, survey respondents often report similar WTP estimates for risks of differing magnitudes (Hammitt and Graham 1999, Corso et al. 2001). One explanation is that respondents may not comprehend the size of these small risk reductions; there is some evidence that visual aids can help improve comprehension and produce WTP estimates that are proportional to the risk reduction (Corso et al. 2001, Hammitt and Haninger 2010). However, even studies that apply these techniques at times find insensitivity (e.g.,. Alberini et al. 2004, Haninger and Hammitt 2011). Proportionality has been identified as a factor that analysts should consider in assessing the quality of stated preference research (e.g., Kling et. al 2011).[12]

Reports in the popular press and the threat of Congressional action may have inhibited EPA's ability to adjust its estimates to reflect new knowledge, particularly if the adjustment is downwards. Controversy first arose over the use of lower values for older individuals, described below. As discussed in Viscusi (2009), the public has tended to become anchored in the current values more generally, and resists decreases in these values in part due to loss aversion. Cameron (2010) notes that this resistance also reflects misunderstanding of the VSL concept. Many appear to interpret the VSL as the value that "the government" is placing on an individual's life, rather than the value that the individuals themselves place on small changes in their own risks. DOT has faced less resistance, possibly because it has consistently proposed increases in the VSL.

3.2.3 Scenario Differences

Perhaps the most difficult issue for those using benefit transfer to value mortality risk reductions has been differences between the study context and the policy context. Until recently, the literature was dominated by studies of job-related risks, while many policy applications address

[12] Tests for sensitivity to risk magnitude generally cannot be implemented in wage-risk studies and thus are not discussed in the DOT (2013) guidance, which does not include stated preference studies.

different types of risks that affect different populations. Economic theory suggests that WTP will vary depending on the characteristics of the individual (such as age and income) and the risk (such as whether it is associated with illness or injury, or results from causes viewed as involuntarily incurred or outside the individual's control). However, without empirical research, the influence of these characteristics on WTP is difficult to predict and the direction and magnitude of the resulting bias is generally unknown.

Traditional approaches to research synthesis, including qualitative systematic review and meta-analysis, may be of limited use in addressing these issues given the lack of data. As introduced earlier, the standard approach for addressing scenario differences is to use a benefit transfer function that relies on data from an individual study or a group of studies to predict the effect of these differences on individual WTP. However, to develop such a function, the relevant parameters (e.g., age, income, morbidity severity and duration) must be included in the underlying studies. Otherwise, the conventional choices include addressing the resulting uncertainties in qualitative discussion, making adjustments based on simple assumptions rather than research results, or supplementing the studies used in the transfer with other research.

One example is age adjustments, which have been a particularly contentious issue for EPA. By definition, the wage-risk studies that dominate the field include only those within the workforce; often between ages 18 and 65. Many of EPA's major regulations address airborne particulate matter and tend to benefit an older population, with some analyses (e.g., EPA 1999) suggesting that roughly 80 percent of the associated mortality risk reductions accrue to individuals over age 65. Prior to 2003, EPA sometimes adjusted for this age difference in sensitivity analysis, using simple assumptions to derive a constant value per statistical life year (VSLY) from the agency's VSL, or adjusting the VSL for age using factors from studies other than those that support the base VSL estimate (see Robinson 2007 for more discussion). Both approaches lead to lower values for older individuals.

While intuition might suggest that it is reasonable to assume that those who are older would report a lower VSL, given that they have fewer remaining years of life expectancy, both theory (Hammitt 2007) and growing empirical evidence suggest that this is not necessarily the case. For example, older individuals may have more accumulated wealth and less interest in spending it on other goods and services, increasing their WTP for reducing their own mortality risks. Some studies (e.g., Aldy and Viscusi 2008) suggest that both the VSL and VSLY rise then

fall over the ages when adults are working, while others suggest that VSL may be constant or decrease slightly after retirement (e.g., Krupnick 2007), which implies VSLY increases with age. As a result of emerging research as well as public concerns, EPA discontinued the practice of adjusting for age, reflecting guidance from OMB and its expert panels (OMB 2003, Cropper et al. 2007, National Academies 2008).[13] Age adjustments appear to be of less concern in DOT analyses, which often address regulations that affect motor vehicle drivers and passengers of all ages. Neither agency currently implements age adjustments.

Another example is the relationship between WTP for deaths that result relatively immediately from injuries and those that result from illnesses preceded by periods of morbidity. When EPA first developed its VSL estimates, most studies considered accidental deaths. As research on WTP to reduce illness-related mortality risks increased, EPA suggested various approaches for adjusting its estimates (EPA 2000b, EPA 2010b), particularly for cancers. However, the research findings are inconsistent and EPA's Science Advisory Board panels suggested that more work was needed before implementing any such adjustments (Stavins 2000, Kling et al. 2011). DOT has not grappled with this issue; its regulations generally address immediate deaths from accidents, similar to job-related risks.

The obvious approach to resolving these issues is to conduct more primary research, but such research requires both time and funding. Moreover, experience to date suggests that future research findings will be mixed, and so many more studies may be required to develop a clearer understanding.

In the interim, what should the analyst do? Using simplifying assumptions not well-supported by theory or empirical evidence raises questions about the validity of the adjustment: the extent to which the adjusted values are more valid than the unadjusted values is unclear. Relying on adjustment factors from other studies can also be problematic. The adjustment may vary across populations and risks, so transferring an adjustment factor across studies involves the same challenges as benefit transfer more generally. The adjustment may also depend on the starting point (e.g., the average age of the sample); applying a constant adjustment calculated at the mean (or median) for different studies may introduce additional uncertainties. Finally, if a study is not of high enough quality nor sufficiently suitable to be used to support the base

[13] Such adjustments were abandoned in part due to public outcry over the use of lower values for older individuals, as noted earlier. See Robinson (2007), Viscusi (2009), and Cameron (2010) for more discussion.

estimates, the rationale for using it to adjust the estimates seems weak. Such adjustments may be best applied in sensitivity analyses to illustrate the potential effects of these factors for decisionmakers. We next discuss the use of various synthesis methods to address these and other types of adjustments.

3.4 Synthesizing Estimates

Over the past 20 years, the number and types of VSL studies have increased dramatically. Viscusi and Aldy (2003) identify over 60 wage-risk studies conducted globally through the year 2000; DOT (2013) identifies an additional 15 U.S. studies published in 2003 or later. The dataset referenced in Lindhjem et al. (2011) includes 74 stated preference surveys from around the world that address diverse populations and risks. The use of meta-analysis to explore variation in the VSL has increased as well; for example, Cropper et al. (2011) identify seven that include U.S. studies, published between 2000 and 2011. Researchers have also begun to experiment with expert elicitation and structural models as alternative or supplementary approaches for combining estimates.

3.4.1 Meta-Analysis

Meta-analysis can be designed to address several goals. For example, it can be viewed as a way to develop a "best" or "central" estimate of the VSL; to explore the extent to which the VSL is affected by variation in the analytic approach and the study context; or to predict how the VSL changes in response to changes in the characteristics of the risk or of those affected. In the first case, the assumption is that there is a "true" underlying VSL that is measured with varying degrees of error in each primary research study. The goal of the meta-analysis is to better characterize this true value by combining the results of several studies and statistically controlling for differences in their design. In the second, there is more explicit consideration of VSL heterogeneity; i.e., the true value is characterized as a function of the study context as well as its design, and the role of the meta-analysis is to better understand the influence of these factors. In the final case, the assumption of a true underlying function is maintained, but the goal is to develop a reduced form model that allows the analyst to predict the VSL in contexts that

differ in certain respects from those studied.[14] For example, the meta-analysis (or more precisely, meta-regression) may be used to predict the VSL based on the age and/or income of those affected.

The available VSL meta-analysis have differing goals and use varying approaches. While all explore sources of variability in the estimates, only some explicitly focus on developing estimates for transfer across countries or policy contexts. They also vary in how they select individual studies and estimates for inclusion, as well as in the statistical methods used (e.g., classical vs. Bayesian) and model specification. The included studies are often dominated by research conducted in the U.S.; wage-risk studies also dominate in cases where the meta-analyses includes only revealed preference or both revealed and stated preference studies.

In Table 1, we summarize meta-analyses completed since 2000 that address U.S. values. Although the estimates are not strictly comparable due to differences in the analyses as well as in when the data were collected (and the effects of income growth over time), they are all within a $2.1 million to $9.3 million range. The estimates currently used by EPA and DOT, discussed above, are at the high end of this range.

[14] See Bergstrom and Taylor (2006) and Nelson and Kennedy (2009) for more detailed and technical discussion of the use of meta-analysis for benefit transfer.

Table 1: Summary of VSL Meta-Analyses

Meta-Analysis	Studies Included	Study Publication Dates	Highlighted VSL Estimate(s)[(a)]	Inflated to 2012 dollars[(b)]
Miller (2000)	68 wage-risk, contingent valuation, and consumer behavior studies from 13 countries	1974-1999	$3.7 million (1995 dollars) for the U.S.	$5.6 million
Mrozek & Taylor (2002)	33 wage-risk studies from the U.S. and other countries	1974-1999	$1.5 million to $2.5 million (1998 dollars)	$2.1 million to $3.5 million
Viscusi & Aldy (2003)	30 U.S. wage-risk studies (separately considers other countries and methods)	1974-2000	$7 million (2000 dollars) for the U.S.	$9.3 million
de Blaeij et al. (2003)	30 road safety consumer behavior, cost per life saved, and contingent valuation studies from the U.S. and other countries	1973-2001	None highlighted	N/A
Kochi et al. (2006)	40 wage-risk and contingent valuation studies from the U.S. and other high income countries	1974-2002	$5.4 million (2000 dollars)	$7.2 million
Bellavance et al. (2009)	37 wage-risk studies from 9 countries.	1974-2004	None highlighted	N/A
Lindhjem et al. (2010); OECD (2012)	74 contingent valuation surveys	1970-2008	$3 million (2005 dollars) for OECD countries	$3.5 million
Dekker et al. (2011)	26 stated preference studies from the U.S. and other countries	1983-2008	$2.4 million to $7.5 million (2004 dollars) in illustrative example	$2.9 million to $9.1 million

Notes:
(a) Estimates highlighted by authors, if any, unless otherwise indicated. Some meta-analyses focus on exploring variation in the estimates rather than highlighting particular values.
(b) Not adjusted for real income growth, which will increase the estimates.

Most of the studies included in the meta-analyses are relatively old, as indicated in the table. These older studies do not necessarily reflect current best practice standards or more recent economic or other conditions that may influence the VSL. Individual U.S. studies that meet current standards for best practices (as described in Kling et al. 2011 and DOT 2013) tend to result in values above the middle of the ranges reported in Table 1. Examples of wage-risk studies include Viscusi (2004) which yields a central estimate of $6.7 million, and Kniesner et al. (2012) which yields a range of $5.2 million to $13.0 million. Both are based on the improved CFOI risk data and better control for factors that influence the relationship between wages and fatality risks than many of the studies included in the meta-analyses. Examples of stated preference studies include Corso et al. (2001) and Hammitt and Haninger (2010) which yield

VSL estimates of $4.1 million to $5.8 million and $6.6 million to $11.1 million respectively. Both demonstrate that the values are sensitive to the size of the risk change, unlike many of the contingent valuation and other stated preference studies included in the meta-analyses summarized in Table 1.

To date, EPA and DOT have used selected meta-analyses to provide point estimates of the VSL, rather than to explore how the VSL might change across contexts.[15] Neither agency currently relies on the available meta-analyses due largely to concerns about their quality.. In particular, EPA's use of meta-analysis has been significantly influenced by its expert advisory panels. Its temporary application of some meta-analytic results in analyses of air pollution regulations stemmed in part from the advice of a Science Advisory Board panel (Cameron et al. 2004) that provided advice on its prospective analysis of the benefits and costs of the Clean Air Act. However, the approaches used in these meta-analyses were criticized in a subsequent review (EPA 2006, Cropper et al. 2007), and EPA's Office of Air and Radiation returned to the practice of relying on the 26 older estimates identified in the EPA guidance.

While the EPA (2006) workgroup noted that meta-analysis is an appropriate technique for estimating the VSL, it recommended against using the then-available analyses. Its concerns included, for example, issues related to: (1) deficient reporting of methods and analytic decisions; (2) differences in how the VSL is defined; (3) criteria for selecting studies and estimates to be included in the analysis; and (4) estimation of the meta-analytic models.[16]

The subsequent Science Advisory Board report (Cropper et al. 2007) recommended a sequence of steps that involve using meta-analysis to better understand how study design and context affect the results. The findings would be used to establish criteria for selecting studies and for selecting estimates from those studies; these estimates could then be combined in a meta-analysis to estimate the VSL. The panel opposed the use of meta-regression to adjust for population or risk characteristics, indicating that meta-analysis should be used only to derive a distribution of values to reflect uncertainty in the estimates. It suggested that adjustments for the policy context should be addressed by selecting an appropriate subset of studies to be applied in each context rather than through meta-regression.

[15] As noted earlier, for a short period, EPA's Office of Air and Radiation used the results of selected meta-analyses to estimate the VSL in their regulatory analyses, while other EPA offices continued to rely on the results from the Viscusi (1992, 1993) review.

[16] Nelson and Kennedy (2009) build on this workgroup's report, examining the use of meta-analysis in environmental economics, discussing how to best address several technical issues, and recommending best practices.

EPA responds to these recommendations in its 2010 White Paper, which was reviewed by another Science Advisory Board panel (Kling et al. 2011). Rather than recommending a particular meta-analytic approach, the panel indicated that the approach should depend on the estimates available. Unlike the prior panel, it suggests that meta-regression may be appropriate for benefit transfer in some cases, and that EPA should also consider other approaches. Perhaps most importantly, it again recommends that EPA recognize the heterogeneity of the VSL and use different sets of studies to estimate the VSL in different contexts to the extent possible.

DOT's 2013 guidance reflects the influence of EPA's work, incorporating some of the conclusions of the 2010-2011 review. DOT weights each of the estimates that underlie its VSL equally, however, rather than using meta-analysis or other statistical techniques to combine them.

The above discussion suggests that there is a role for meta-analysis in estimating the VSL, but that more work is needed to develop analyses suitable for use in regulatory assessments. In addition, EPA's expert panels appear somewhat optimistic about the ability to use different sets of studies to address different types of risks. Reviews of the available research suggest that this may be difficult in the absence of substantial new primary research. Many wage-risk and stated preference studies include basic population characteristics (such as age or income) as explanatory variables, so can be used to address the effects of these characteristics on the VSL, although the results still remain somewhat uncertain.[17] Fewer studies address variation in other population characteristics, such as underlying health status, or in risk characteristics. Values may vary, for example, depending on the severity and duration of illness prior to death, on the characteristics of the risk-reducing program, and on whether the risk is perceived as ambiguous, particularly dreaded, involuntarily, or uncontrollable (see Robinson et al. 2010).

Addressing how these risk characteristics affect the VSL generally requires the use of stated preference research that can be tailored to the outcomes of concern. However, it is unclear whether sufficient high quality studies are available to support estimates that specifically address environmental or other risks. For example, the Lindhjem et al. (2011) database (described in more detail in OECD 2012), includes 74 stated preference surveys from around the world, of which 37 address traffic safety, 30 address public health or health risks from unidentified causes,

[17] See earlier discussion on the effects of age; Hammitt and Robinson (2010) summarize research on the effects of income.

and 17 address environmental risks (some address more than one risk type). Of the 17 that address environmental risks, many may not be suitable for use in EPA regulatory analysis, either because they address dissimilar risks or populations or because of concerns about their quality. Thus even as approaches to VSL meta-analysis improve as a result of the above and other efforts, some benefit transfer issues will remain difficult to address.

3.4.2 Structural Models

Another approach to estimating the VSL is structural modeling. Such modeling is described as a promising approach in the EPA Science Advisory Board review discussed above (Kling et al. 2011), although the panel notes that the approach needs to be further developed before it will be suitable for use in regulatory analysis.

Structural benefit models (or "preference calibration") (Smith et al. 2002) involve estimating a series of functions that relate an individual's VSL to characteristics such as their age and income. The functions are based on parametric utility functions and are calibrated to be consistent with available estimates for various benefits.[18] As the functions are derived from an underlying utility function, the estimates are consistent with theoretical restrictions; for example, that estimates of WTP do not exceed individual income and are responsive to changes in the size of the risk reduction. Data drawn from several sources are used to estimate these functions, which can be applied to predict the variation in individual WTP attributable to changes in the modeled parameters. The advantages of such models are that they combine data from different studies that do not necessarily address the same outcome, and ensure both internal consistency and consistency with theory. Thus they are less reliant on completion of studies that use the same outcome measure and can be used to predict values in contexts outside of those studied. The disadvantage is that they rely on analysts' ability to develop an appropriate valuation function and estimate its parameters.

Smith et al. (2006) illustrate this approach by combining data from VSL wage-risk and stated preference studies as well as other sources such as the Health and Retirement Survey and the National Income Product Accounts. They develop a series of functions that relate individual willingness to trade-off job-related income for reduced mortality risk to decisions about whether

[18] The term "utility" is used by economists to refer to individual well-being as self-defined.

and how much to engage in paid labor. They use their approach to show how the variation in baseline mortality risk with age affects the values individuals place on changes in job-related risks at each age. They report estimated VSLs ranging from about $6 million to $10 million (1983 dollars), or $13.8 million to $23.1 million if inflated to 2012 dollars, depending on the model specification and age group affected.

In another example, Newbold (2011) extends this approach to a life-cycle consumption model that relates individual utility (well-being) to changes in health and income over the life-span. This approach can be used to integrate valuation of morbidity and mortality risks as well as to consider the characteristics of those affected. At this point, it is primarily conceptual. No empirical estimates are provided.

Such structural models could be very useful in conducting benefit transfer, but require additional work to further develop the conceptual framework and to use that framework to develop empirical estimates. They have the potential to provide values that go beyond the relationships that are currently estimated in empirical studies or that can be estimated in meta-regressions, by relying on theoretical models and supplementary data sources. However, the extent to which they can accommodate the diverse characteristics that affect the VSL is unclear, particularly where theoretical expectations (and empirical data) are ambiguous.

3.4.3 Expert Elicitation

Another approach is structured expert elicitation (EPA 2011). It is a formal approach, often used in risk analysis, that provides a systematic framework for obtaining experts' judgments about the value of specific parameters. Each expert is asked to provide his or her judgments in the form of a probability distribution, which characterizes his or her beliefs about the value of the parameter (e.g., whether it is equally likely to be larger or smaller than the median of the expert's distribution).

Expert elicitation requires the experts to be transparent about their judgments; in addition, experts are often asked to explain their use of evidence to support their judgments. It is identified as an approach that could inform the analysis in the Science Advisory Board's 2007 and 2011 VSL reviews, although the earlier panel indicates that meta-analysis is preferred when possible.

The expert elicitation process is designed to avoid well-known heuristics and biases that can lead to poor judgment. It begins with developing a well-defined question to be addressed and

an extensive elicitation protocol that is pre-tested. The protocol is designed to follow a logical structure that ensures that the experts are each answering the same questions and clearly explain the basis for their responses. Experts are identified through a formal process, which may include review of publications and peer nomination, intended to provide an unbiased range of perspectives. Those conducting the elicitation often provide the experts with background materials and hold a pre-elicitation workshop. The elicitation is conducted with each expert individually, often through a lengthy interview following the pre-determined protocol.

Expert elicitation has been used to estimate the VSL in one pilot study (Roman et al. 2012) that focuses on air pollution risks. Three experts were first asked to quantify the VSL for a base case defined as accidental death to a 40 year old Canadian of average income, then to quantify adjustments for application to deaths from air pollution-related illnesses of Canadians of different age, income, and health status. The experts' median VSL estimates for the base case range from about CAN 6 million to CAN 12 million (2007 Canadian dollars) (i.e., $7.4 million to $14.8 million in 2012 US dollars), and their interquartile ranges (i.e., central 50 percent credibility intervals) jointly span the range from about $3.5 million to $13 million ($4.3 million to $16.1 million).

These values are comparable to the estimates from the meta-analyses discussed above, which is not surprising as the experts relied on many of the same primary studies. However, the authors report that each of the experts expressed reservations about the meta-analyses (especially their inclusion of older studies using inferior data or methods) and instead based their judgments primarily on selected primary studies. Combining base case VSL with adjustment factors for age (to 60 years old) and death from air pollution-related illness led to some convergence among the experts, with median VSLs of CAN 5 million to CAN 7.5 million ($6.2 million to $9.3 million) and interquartile ranges spanning the range from CAN 4.5 million to CAN 13.5 million ($5.6 million to $16.7 million). None of the meta-analyses address the question of extrapolating to other ages or to air pollution-related illness, so we cannot provide any comparative results. Roman et al. conclude that the pilot provided useful insights about the VSL and suggest improvements in the elicitation process; this pilot has not yet been expanded to a full-scale elicitation.

Like the other approaches discussed in this paper, expert elicitation cannot be conducted in a vacuum; both theory and empirical research are needed to provide the foundation for the

experts' judgments. However, it provides an option for extrapolating from the contexts studied to other contexts, offering insights into the direction and magnitude of the changes in the VSL and related uncertainties.

4.0 CONCLUSIONS AND IMPLICATIONS

The discussion in the preceding sections suggests that there is not a "one-size-fits-all" approach to synthesizing research for benefit transfer. The approach will depend on the quality of the available research as well as the match between the study and policy context. The VSL literature provides several examples. In a case where the outcome has been well-studied, such as job-related risks, it may be relatively straightforward to select high quality studies and use established meta-analytic statistical techniques to synthesize the results. As the match between the available high quality research and the policy context weakens, the analyst faces more difficult decisions. Because some demographic characteristics, such as age and income, are included in many of the primary research studies, understanding their effect may be straightforward as long as risk characteristics are held constant. However, because the effects of age and income may vary as risk characteristics change, estimating these effects across contexts poses greater challenges. In addition, some risk characteristics are not well-studied or consistently defined across studies, and theory does not provide clear guidance about their likely impact. Thus, not surprisingly, the approach for tailoring the VSL to a particular context becomes more speculative as the applicable research base becomes less well-populated.

The approaches that are most well-developed and widely-used in regulatory analysis include applying point estimates from individual studies or ranges of point estimates from a group of selected studies. As a result of EPA's substantial investment in exploring the research literature and alternative approaches to synthesizing this research, as well as in review by independent advisory panels, we now have a much better sense of the potential role of meta-analysis in improving our understanding of the literature − but the results of applying these recommendations remain to be seen. Comparison of the older meta-analyses to newer primary research studies suggest that improved data and practices may yield somewhat higher U.S. estimates.

Qualitative literature reviews or meta-analysis are less useful in addressing the effects of variation in population and risk characteristics that have not been well-studied. Both structural

models and expert elicitation appear more promising for extrapolating from the contexts studied to other contexts. In the case of structural models, additional work is needed to better develop the conceptual framework and use it for empirical estimation. However, structural models are primarily useful where there are clear theoretical expectations. It is not clear how structural models could help in adjusting the VSL to different types of mortality risks (e.g., air pollution-related disease), where theory is ambiguous. Expert elicitation is more flexible as it does not require the existence of a well-developed research data base. Full-scale implementation of an expert elicitation in this context would be informative. Each approach relies to varying extents on analysts' judgments and statistical tools. A key question is how to ensure that this judgment is well-informed and that the rationale is clearly articulated along with the analytic results.

Throughout this literature, inconsistent reporting of study methods, assumptions, and results inhibits the use of synthesis methods and the application of the results in regulatory analysis. Establishment of reporting standards that are acceptable to researchers working in this field is clearly needed. The quality of any benefit transfer relies largely on the availability of research that applies best practice standards. Further articulation of these standards and enhanced funding of primary research that meets such standards would be helpful.

Regardless of whether primary research or benefit transfer is used to estimate the value of nonmarket outcomes such as reduced mortality risks, the accuracy of the results will remain uncertain. Some uncertainties result from the need to make judgments and assumptions in any research design; others relate to the match between the study and policy context. To be useful for decisionmaking, the application of the available evidence must be accompanied by discussion of its limitations and the implications for related policy decisions. In addition, the extent to which investment in further analysis is worthwhile will depend on the policy context. At times, refining benefit values may be of major importance; in other cases, decisionmakers may be better served by further investment in other analytic components.

REFERENCES

Alberini, A., M. Cropper, A. Krupnick, and N. 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" *Review of Economics and Statistics.* 90(3), 573–581.

Ashenfelter, O. and M. Greenstone. 2004. "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *The American Economic Review* 94(2):454–460.

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(2): 444–464.

Bergstrom, J.C. and L.O. Taylor. 2006. "Using Meta-Analysis for Benefits Transfer: Theory and Practice." *Ecological Economics*. 60: 351-360.

Brouwer, R. 2000. "Environmental Value Transfer: State of the Art and Future Prospects. *Ecological Economics.* 32(1): 137-152.

Cameron, T.A. et al. 2004. *Review of the Revised Analytical Plan for EPA's Second Prospective Analysis – Benefits and Costs of the Clean Air Act 1990 – 2020.* An Advisory by a Special Panel of the Advisory Council on Clean Air Compliance Analysis. EPA-SAB-COUNCIL-ADV-04-004.

Cameron, T.A. 2010. "Euthanizing the Value of a Statistical Life." *Review of Environmental Economics and Policy.* 4(2): 161-178.

Clinton, W.J. 1993. "Executive Order 12866: Regulatory Planning and Review." *Federal Register*. 58(190): 51735-51744.

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.

Cropper, M. et al. 1999. "Final Advisory by the Advisory Council on Clean Air Compliance Analysis on the 1999 Prospective Study of Costs and Benefits of Implementation of the Clean Air Act Amendments." Memorandum from the Chair, Advisory Council on Clean Air Act Compliance Analysis, to EPA Administrator Christine Todd Whitman.

Cropper, M. et al. 2001. "Arsenic Rule Benefits Analysis; An EPA Science Advisory Board Review." Memorandum from the Chair, Science Advisory Board, and the Chair, Arsenic Rule Benefits Review Panel, to EPA Administrator Christine Todd Whitman. EPA-SAB-EC-01-008.

Cropper, M. et al. 2007. "SAB Advisory on EPA's Issues in Valuing Mortality Risk Reduction." Memorandum from the Chair, Science Advisory Board, and the Chair, Environmental Economics Advisory Committee, to EPA Administrator Stephen L. Johnson. EPA-SAB-08-001.

Cropper, M., J.K. Hammitt, and L.A. Robinson. 2011. "Valuing Mortality Risk Reductions: Progress and Challenges." *Annual Review of Resource Economics.* 3: 313-336.

de Blaeij, A. et al. 2003. "The Value of Statistical Life in Road Safety: A Meta-Analysis." *Accident Analysis & Prevention.* 35(6): 973-986.

Dekker, T., R. Brouwer, M. Hofkes, and K. Moeltner. 2011. "The Effect of Risk Context on the Value of a Statistical Life: A Bayesian Meta-Model." *Environmental and Resource Economics.* 49(4): 597-624.

Dockins, C., K. Maguire, N. Simon, and M. Sullivan. 2004. *Value of Statistical Life Analysis and Environmental Policy: A White Paper.* Prepared by the National Center for Environmental Economics, U.S. Environmental Protection Agency, for Presentation to the Environmental Economics Advisory Committee of the Science Advisory Board.

Dockins, C., K. Maguire, and N. Simon. 2006. *Willingness to Pay for Environmental Health Risk Reductions When There are Varying Degrees of Life Expectancy: A White Paper.* Prepared by the National Center for Environmental Economics, U.S. Environmental Protection Agency, for a Consultation with the Environmental Economics Advisory Committee of the Science Advisory Board.

Doucouliagos, C., T. D. Stanley, and M. Giles. 2012. "Are Estimates of the Value of a Statistical Life Exaggerated?" *Journal of Health Economics.* 31(1): 197-206.

Hammitt, J.K. 2000. "Valuing Mortality Risk: Theory and Practice." *Environmental Science and Technology*. 34: 1396-1400.

Hammitt, J.K. 2007. "Valuing Changes in Mortality Risk: Lives Saved versus Life Years Saved." *Review of Environmental Economics and Policy*. 1(2): 228-240.

Hammitt, J.K. et al. 2010. "Advisory on a Preliminary Draft of the Second Section 812 Prospective Study of the Benefits and Costs of the Clean Air Act (April 2010)." Memorandum from the Chair, Advisory Council on Clean Air Compliance Analysis, to EPA Administrator L.P. Jackson. EPA-COUNCIL-10-004.

Hammitt, J.K. and J.D. Graham. 1999. "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" *Journal of Risk and Uncertainty.* 18(1): 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.

Hammitt, J.K. and L.A. Robinson. 2011. "The Income Elasticity of the Value per Statistical Life: Transferring Estimates Between High and Low Income Populations." *Journal of Benefit-Cost Analysis.* 2: Art. 1.

Haninger, K. and J.K. Hammitt. 2011. "Diminishing Willingness to Pay per Quality-Adjusted Life Year: Valuing Acute Foodborne Illness." *Risk Analysis.* 31: 1363-1380.

Industrial Economics, Incorporated. 2003. *Benefits and Costs of the Clean Air Act 1990 - 2020: Revised Analytical Plan For EPA's Second Prospective Analysis.* Prepared for J. DeMocker, U.S. Environmental Protection Agency.

Johnston, R.J., E.Y. Besedin, and M.H. Ranson. 2006. "Characterizing the Effects of Valuation Methodology in Function-Based Benefits Transfer." *Ecological Economics*. 60(2): 407-419.

Kaul, S. et al. 2013. "What Can We Learn from Benefit Transfer Errors? Evidence from 20 Years of Research on Convergent Validity." *Journal of Environmental Economics and Management*. 66: 90-104.

Kling, C.L. et al. 2011. "Review of 'Valuing Mortality Risk Reductions for Environmental Policy: A White Paper' (December 10, 2010)." Memorandum to Lisa P. Jackson, EPA Administrator, from the EPA Science Advisory Board and Environmental Economics Advisory Committee. EPA-SAB-11-011.

Kniesner, T. J., W.K. Viscusi, C. Woock, and J.P. Ziliak. 2012. "The Value of Statistical Life: Evidence from Panel Data." *Review of Economics and Statistics.* 94:19-44.

Kochi, I., B. Hubbell and R. Kramer. 2006. "An Empirical Bayes Approach to Combining and Comparing Estimates of the Value of a Statistical Life for Environmental Policy Analysis." *Environmental and Resource Economics.* 34: 385-406.

Krupnick, A. 2007. "Mortality-risk Valuation and Age: Stated Preference Evidence." *Review of Environmental Economics and Policy.* 1(2): 261-282.

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(9): 1381-1407.

Miller, T.R. 1990. "The Plausible Range for the Value of Life – Red Herrings Among the Mackerel." *Journal of Forensic Economics*. 3(3): 17-40.

Miller, T.R. 2000. "Variations Between Countries in Values of Statistical Life." *Journal of Transport Economics and Safety*. 34: 169-188.

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.

National Academies. 2008. *Estimating Mortality Risk Reduction and Economic Benefits from Controlling Ozone Air Pollution*. Committee on Estimating Mortality Risk Reduction Benefits from Decreasing Tropospheric Ozone Exposure. Washington D.C.: National Academies Press

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.

Newbold, S.C. 2011. "Valuing Health Risk Changes Using a Life-Cycle Consumption Framework." National Center for Environmental Economics, U.S. Environmental Protection Agency, Working Paper 11-03.

Obama, B. 2011. "Executive Order 13563: Improving Regulation and Regulatory Review." *Federal Register*. 76(14): 3821-3823.

Organisation for Economic Co-operation and Development. 2012. *Mortality Risk Valuation in Environment, Health and Transport Policies.* Paris: OECD Publishing.

Ready, R. et al. 2004. "Benefit Transfer in Europe: How Reliable are Transfers between Countries?" *Environmental and Resource Economics.* 29: 67-82.

Robinson, L.A. 2007. "How U.S. Government Agencies Value Mortality Risk Reductions." *Review of Environmental Economics and Policy.* 1(2): 283-299.

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

Robinson, L.A., J.K. Hammitt, J.E. Aldy, A. Krupnick, and J. Baxter. 2010. "Valuing the Risk of Death from Terrorist Attacks." *Journal of Homeland Security and Emergency Management.* 7(1): Art. 14.

Robinson, L.A. and J.K. Hammitt. 2013. "Skills of the Trade: Valuing Health Risk Reductions in Benefit-Cost Analysis." *Journal of Benefit-Cost Analysis*. 4(1): 107-130.

Roman, H.A. J.K. Hammitt, T.L. Walsh, and D.M. Stieb. 2012. "Expert Elicitation of the Value Per Statistical Life in an Air Pollution Context." *Risk Analysis.* 32: 2133-2151.

Rosenberger, R.S. and J.B. Loomis 2003. "Benefit Transfer." *A Primer on Non-Market Valuation* (Champ, P.A., K.J. Boyle, and T.C. Brown, eds.). Boston MA: Kluwer Academic Press.

Rosenberger, R.S. and T.D. Stanley. 2006. "Measurement, Generalization, and Publication: Sources of Error in Benefit Transfers and their Management." *Ecological Economics.* 60: 372-378.

Schmalensee, R. et al. 1993. "Science Advisory Board's Review of the Office of Policy, Planning, and Evaluation's (OPPE) and Office of Air and Radiation's (OAR) Progress on the Retrospective and Prospective Studies of the Impacts of the Clean Air Act." Letter from the Chair, Science Advisory Board's Clean Air Act Compliance Analysis Council, to EPA Administrator Carol M. Browner.

Smith, V.K., S.K. Pattanayak, and G.L. van Houtven. 2006. "Structural Benefits Transfer: An Example Using VSL Estimates." *Ecological Economics.* 60: 361-371.

Smith, V.K., G. Van Houtven, and S.K. Pattanayak. 2002. "Benefits Transfer Via Preference Calibration: "Prudential Algebra" For Policy." *Land Economics.* 78(1): 132–152.

Stavins, R. et al. 1999. "An SAB Report on the EPA *Guidelines for Preparing Economic Analyses*." Memorandum from the Chair, Science Advisory Board, and the Chair, Environmental Economics Advisory Committee, to EPA Administrator Carol M. Browner. EPA-SAB-EEAC-99-020.

Stavins, R. et al. 2000. "An SAB Report on EPA's White Paper *Valuing the Benefits of Fatal Cancer Risk Reduction*." Memorandum from the Chair, Science Advisory Board, and the Chair, Environmental Economics Advisory Committee, to EPA Administrator Carol M. Browner. EPA-SAB-EEAC-00-013.

U.S. Department of Transportation. 1993. *Treatment of Value of Life and Injuries in Preparing Economics Evaluations.* Memorandum to Assistant Secretaries and Modal Administrators from W.B. McCormick Jr., General Counsel, and J.N. Shane, Assistant Secretary for Policy and International Affairs.

U.S. Department of Transportation. 2008. *Treatment of the Economic Value of a Statistical Life in Departmental Analyses*. Memorandum to Secretarial Officers and Modal Administrators from T.D. Duvall, Assistant Secretary for Transportation Policy, and D.J. Gribbin, General Counsel.

U.S. Department of Transportation. 2013. *Guidance on Treatment of the Economic Value of a Statistical Life (VSL) in Departmental Analyses*. Memorandum to Secretarial Officers and Modal Administrators from P. Trottenberg, Under Secretary for Policy, and R. Rivkin, General Counsel.

U.S. Environmental Protection Agency. 1997. *The Benefits and Costs of the Clean Air Act, 1970 to 1990.*

U.S. Environmental Protection Agency. 2000a. *Guidelines for Preparing Economic Analysis*. EPA 240-R-003.

U.S. Environmental Protection Agency. 2000b. *Valuing Fatal Cancer Risk Reductions*.

U.S. Environmental Protection Agency. 2006. *Report of the EPA Work Group on VSL Meta-Analyses*.

U.S. Environmental Protection Agency. 2010a. *Guidelines for Preparing Economic Analysis*. EPA 240-R-10-001.

U.S. Environmental Protection Agency. 2010b. *Valuing Mortality Risk Reductions for Environmental Policy: A White Paper (Review Draft)*. Prepared by the National Center for Environmental Economics for consultation with the Science Advisory Board – Environmental Economics Advisory Committee.

U.S. Environmental Protection Agency. 2011. *Expert Elicitation Task Force White Paper.* Prepared for the Science and Technology Policy Council.

U.S. Office of Management and Budget. 2003. *Circular A-4: Regulatory Analysis.*

U.S. Office of Management and Budget. 2013. *2013 Draft Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance with the Unfunded Mandates Reform Act.*

Urban Institute. 1991. *The Costs of Highway Crashes, Final Report*.

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

Viscusi, W.K., 1993. "The Value of Risks to Life and Health." *Journal of Economic Literature.* 31: 1912-1946.

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

Viscusi, W.K. 2009. "The Devaluation of Life." *Regulation and Governance.* 3:103-127.

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