From Minor Ailments to Major Illnesses: A Practical Framework for Valuing Morbidity in Environmental Policy Analysis[1]

Mark Dickie
University of Central Florida

Matthew Neidell
Columbia University, NBER and IZA

Abstract: Environmental decisions often impact health, yet morbidity valuation remains challenging, with current approaches like cost-of-illness (COI) measures and Quality-Adjusted Life Years failing to reflect individuals' willingness to pay to avoid adverse health states. This paper introduces a novel framework for valuing morbidity across health conditions, deriving willingness-to-pay (WTP) estimates using widely available preference weights and a baseline condition with known WTP. Our analysis reveals WTP estimates significantly exceeding commonly used COI measures, indicating undervaluation of health states in policy evaluations. Sensitivity analyses highlight variations consistent with those for the value of statistical life, underscoring the method's robustness. This framework offers a practical tool for integrating morbidity into analyses of environmental policy.

[1] This paper emerged from discussions at the Environmental Protection Agency's Science Advisory Board meeting on the BenMAP & Benefits Methods Panel. The authors thank Pete Maniloff, Jim Hammitt, and Kristian Bolin for helpful comments.

1. Introduction

Protection of human health is one of the main motivations for environmental regulation, underscoring the importance of valuing health, both mortality and morbidity, in comprehensive policy analysis. While significant strides have been made in valuing mortality, a new approach is needed for morbidity.[2]

Economists have made considerable advances in valuing mortality through the development of the Value of Statistical Life (VSL). The derivation of an estimable expression for VSL based on utility maximization principles has led to a rich empirical literature that has produced reliable and robust VSL estimates across various contexts. This framework allows for the incorporation of mortality risks into environmental policy analysis, with its widespread adoption indicating its importance in decision-making processes. Indeed, the value of reduced premature mortality comprises the majority of monetized benefits of most air pollution regulations (Cropper, Joiner and Krupnick 2024).

Valuing morbidity remains more challenging than mortality valuation, largely due to the vast array of distinct morbidity conditions impacted by environmental regulations. These conditions span a wide spectrum, from relatively minor and transient illnesses to severe, chronic, and life-altering diseases. While Quality-Adjusted Life Years (QALYs), which combine information on life expectancy and health-related quality of life into a single index, estimates exist for

[2] The Office of Management and Budget's updated Circular A-4 (issued in 2023 and rescinded in 2024), which guides federal agencies on Regulatory Impact Analysis, emphasizes the importance of and challenges in valuing nonfatal health effects. Similarly, the recent "Non-fatal Health Effects Workshop," held as part of the Frontier of Cost-Benefit Analysis and organized by the Council of Economic Advisors, Office of Information and Regulatory Affairs, and Office of Science and Technology Policy, underscores these same issues.

various conditions, they lack direct applicability to cost-benefit analysis and face criticism for theoretical inconsistency (Hammitt 2013).[3] Consequently, policy evaluations often rely on cost of illness measures, which estimate direct medical expenditures and lost productivity but omit broader welfare losses, and are therefore widely viewed to be incomplete and likely represent a lower bound of true health value (Harrington and Portney, 1987). This dependence on imperfect measures highlights the ongoing struggle to accurately capture the full economic impact of morbidity in policy assessments.

This paper addresses these longstanding challenges by introducing a novel method for valuing morbidity in a theoretically consistent and practical, policy-useful framework. Our approach is grounded in a standard expected utility model that considers two distinct health conditions. From this model, we derive a straightforward equation that expresses the willingness to pay (WTP) for one condition as a function of another, which we designate as the baseline condition. The relationship between these conditions is scaled by their respective preference weights, reflecting the relative impact on individual well-being. This approach can be viewed as a form of benefit transfer, using a well-characterized baseline condition to infer willingness to pay for other health states.

Applying this method to value numerous health conditions requires three inputs: preference weights for each condition and a known WTP to avoid the baseline condition. For obtaining preference weights, we rely on utility weights,[4] a crucial component of QALYs. Unlike

[3] A primary reason for the theoretical inconsistency of QALYs is that they assume that the utility of a health state is independent of the duration for which that state is experienced.
[4] Preference weights are also commonly referred to as utility weights or Health-Related Quality of Life (HRQoL) weights.

QALYs, utility weights are theoretically sound, having been derived from stated preference elicitation methods that offer a valuation of health conditions on a 0-1 scale, such as Standard Gamble (SG), which elicits choices under uncertainty; Time Trade-Off (TTO), which asks respondents to trade longevity for health quality; and Visual Analog Scale (VAS), which asks respondents to rate health states on a scale.[5] Since these utility weights are a component of QALYs, they are available for thousands of health states, enabling us to apply this method to value morbidity for numerous conditions once we have a reliable WTP to avoid a baseline condition.

We apply this method to several well-defined health conditions, choosing chronic kidney disease (CKD) as the baseline condition because of a recent, comprehensive analysis of its WTP. We obtain utility weights (averaged across elicitation method) from two analyses that provide estimates for multiple health states. Using a mean WTP of $0.68 million to avoid CKD as the baseline (we express all monetary values in 2022$, consistent with the baseline estimate for CKD), we calculate estimates of $0.43 million for mild angina (95% confidence interval 0.11m - 0.74m) and $1.59 million for a major stroke (95% CI 1.17m – 2.05m), to name two. These WTP estimates for mild angina and major stroke far surpass the COI estimates ($49,000 and $34,000, respectively) used in the EPA's Benefits Mapping and Analysis Program (BenMAP), a crucial tool for environmental policy decisions. Our analysis reveals that WTP estimates significantly exceed commonly used COI measures, suggesting an undervaluation of health states in current policy evaluations, likely leading to suboptimal policy implementation.

[5] Although the EQ-5D is a commonly used measure for utility weights, we avoid this measure because of insufficient variation (described in more detail below).

Utility weights have been criticized because different elicitation methods yield different values; we treat this as a potential measurement error issue and address it by producing estimates separately by elicitation method. We focused on VAS and TTO methods, as these were available for both the baseline condition of CKD and other health states being valued. We estimate morbidity values for chronic bronchitis at $0.47 million (VAS) and $0.78 million (TTO), and for major depression at $1.2 million (VAS) and $2.6 million (TTO). While these results indicate variation depending on elicitation method, they consistently rank conditions by method. To put these values in perspective, a comprehensive review of VSL estimates indicates a typical range from approximately $7-$15 million (Viscusi and Aldy, 2003, updated from 2000 U.S. dollars using the Consumer Price Index for All Urban Consumers), which exceeds the ranges we obtain for morbidity, consistent with the expectation that values for nonfatal health risks are smaller than those for mortality risk, though the values remain of a similar order of magnitude.

In addition to applying the method to estimate WTP, we use data from a meta-analysis of morbidity valuation research to test two key theoretical restrictions on the relationship of WTP with preference weights and duration of morbidity. The restriction implied by the model underlying our method is supported by the data, but the additional restriction needed for QALYs to adequately represent preferences is rejected.

The primary objective of this study is not to produce an exhaustive list of morbidity valuations, but rather to demonstrate a practical, "shovel ready" method that can be readily applied to environmental and other policies involving morbidity assessments. We acknowledge that the range in our estimates is broader than ideal, though some degree of uncertainty is inherent in such valuations and the range is not unlike that for VSL. These varying estimates can

be effectively utilized to generate a range of morbidity values as part of a comprehensive sensitivity analysis, which is a crucial component of any robust policy evaluation. By presenting this flexible and adaptable framework, we aim to provide a valuable tool for policymakers and researchers to more accurately incorporate morbidity valuations into their analyses, ultimately leading to more informed and effective policy decisions.

While we see value in our approach, we view it as a first step toward improving the valuation of morbidity. There are several important features that warrant further refinement. Empirically, enhancing the application of utility weights and exploring alternative baseline health conditions could improve the accuracy and robustness of the estimates. Theoretically, our framework focuses on valuing the prevention (i.e., risk reduction) of morbidity, assumes no monetary losses associated with illness, and anchors utility losses to a benchmark of ideal health. In the online Appendix, we outline how the approach could be extended to incorporate these and other dimensions. We acknowledge that additional limitations likely remain, and we hope this work spurs further efforts to develop and refine this approach for valuing morbidity.

Our paper is organized as follows. The next section introduces the theoretical framework for our analysis. The subsequent section outlines our methods, including parameter selection, approaches for assessing measurement error, and validity checks on the model. The results are then presented, accompanied by sensitivity analyses, before concluding with a summary of findings and implications.

2. Theory

We specify a model to express the marginal willingness to pay (WTP) to reduce risk of one non-fatal health condition as a function of (1) the WTP to reduce risk of another non-fatal

condition and (2) the preference weights for the two conditions. The online Appendix outlines how to adapt the model to measure WTP for treatment rather than prevention of morbidity. In the one-period model, an individual may be healthy (H) or ill/injured with one of two conditions (S_1, S_2). The von Neumann-Morgenstern utility function (*U)* depends on income (I) and health state. Expected utility equals

$$EU = p_1 U\left(I, S_1\right) + p_2 U\left(I, S_2\right) + \left(1 - p_1 - p_2\right) U(I, H), \quad (1)$$

where p_1 and p_2 denote the probabilities of the two conditions. Treating the conditions as mutually exclusive aligns with standard morbidity valuation practices, where each outcome is valued independently. For chronic conditions, $U\left(I, S_j\right)$ reflects spending the remaining lifespan with the impairment while for acute conditions, $U\left(I, S_j\right)$ reflects being ill for the relevant period and then recovering (Hammitt 2017). Assume utility increases with income, and for any income level, utility of each illness is less than utility if healthy, $U\left(I, S_j\right) < U(I, H), \; j = 1,2.$

Marginal WTP to reduce each morbidity risk (W_j) monetizes the utility improvement from being healthy over being sick:

$$W_j = \left[U(I,H) - U\left(I, S_j\right)\right] \Big/ MEUI, \qquad j = 1,2, \quad (2)$$

where the denominator represents the marginal expected utility of income.[6] Define the preference weight q_j as utility with condition *j* relative to utility if healthy:

$$q_j = U\left(I, S_j\right) \Big/ U(I, H), \quad j = 1, 2. \quad (3)$$

[6] Differentiate equation (1) with respect to morbidity risk and income and apply the implicit function theorem and rule to derive equation (2). Also,
$MEUI = p_1\left(\partial U\left(I, S_1\right) \Big/ \partial I\right) + p_2\left(\partial U\left(I, S_2\right) \Big/ \partial I\right) + \left(1 - p_1 - p_2\right)\left(\partial U(I,H) \Big/ \partial I\right).$

Substituting q_j into the expressions for W_j in equation (2) and solving yields:

$$W_1 = \frac{1-q_1}{1-q_2} W_2, \tag{4}$$

expressing WTP to reduce one morbidity risk as the WTP to reduce another, rescaled by their relative disutilities.

Equation (4) forms the basis of our method. For consistency with the one-period model and because WTP varies with illness duration whereas preference weights do not, equation (4) would apply either to two chronic conditions or to two acute conditions.[7] Typically, WTP to avoid the baseline condition (W_2) is estimated by contingent valuation or choice experiment, while preference weights are measured by techniques like SG, TTO or VAS (Torrance 1985, Institute of Medicine 2006).[8]

Our approach offers timely estimation of WTP to avoid nonfatal health outcomes, leveraging existing preference weights for many conditions. Primary valuation studies, although ideal, would take years and require substantial financial resources to address the diverse

[7] Other differences between preference weight and WTP measures are that weights usually assume no monetary losses from morbidity and measure utility of illness relative to utility in ideal health, whereas WTP measures often account for monetary losses and compare utility of illness to utility with the actual health state, which may be less than ideal. Effects of these differences on the WTP measure in equation (4) are examined in the online Appendix.

[8] Usually, preference weights are scaled so that utility =1 if perfectly healthy and =0 if dead. Each measurement method can be adapted to represent health states considered worse than death using negative values. Measurement of preference weights invokes assumptions beyond those needed to derive equation (4), including independence of preference weights from income and duration of morbidity (see Bleichrodt and Quiggin 1999, Miyamoto et al. 1998, Hammitt 2013).

morbidity outcomes relevant to policy, and illness costs provide only a lower bound on WTP. Meanwhile, monetized quality-adjusted life years (QALYs) require further assumptions for compatibility with benefit-cost analysis and likely are inconsistent with individual preferences (Hammitt 2013).[9]

Kenkel (2006) proposed a related approach, expressing WTP to reduce morbidity risk as a fraction of WTP to reduce mortality risk (denoted W_r) : $W_1 = w\left(1-q_1\right)W_r$, where w is close to 1.[10] Taking mortality risk as the baseline exploits the health outcome with the most robust valuation evidence. Conversely, our approach aligns with benefit transfer recommendations concerning similarity between baseline and valued outcomes (HHS 2016 Johnston et al. 2021, OMB 2024, Rosenberger and Loomis 2017), as willingness to pay depends on characteristics such as severity, duration, and the nature of the risk, and transferring values across dissimilar contexts can introduce systematic bias, whereas greater similarity improves the validity of transferred estimates.[11] Two nonfatal conditions share common features distinct from mortality (continuation of life, availability of treatment).[12]

Additionally, our approach can incorporate multiple baseline conditions when available, enabling assessment of validity or improved WTP estimation. As suggested by equation (4),

[9] QALYs are a nonmonetary measure that integrates severity and duration of health states into a scalar. When lacking credible WTP estimates, some federal agencies value morbidity using a monetary value per QALY, often based on the value of a statistical life year (HHS 2016).
[10] Kenkel's model includes mortality risk and a single morbidity risk.
[11] Benefit transfer refers to using a WTP estimate or function for one outcome to value another outcome.
[12] Our method also may be less sensitive to different methods of measuring preference weights. Researchers often report that standard gamble weights exceed time-tradeoff weights which exceed visual-analog-scale weights. If the same method is used to measure q_1 and q_2, the impact of any tendency of the method to over- or under-state the underlying preference may be mitigated by using the ratio in equation (4).

expanding the model to include several nonfatal conditions implies that the ratio $W_j \big/ \left(1 - q_j\right)$ is equal for all conditions, a testable restriction.[13] If there are *M* conditions, then $W_j \big/ \left(1 - q_j\right) = e^{\alpha},\ j = 1,...,M,$ where α is a constant and

$$(5) \qquad \ln W_j = \alpha + \beta_1 \ln\left(1 - q_j\right), \quad \beta_1 = 1, \quad j = 1,...,M.$$

If the conditions differ in duration, the restriction would be tested after removing effects of duration:

$$(6) \qquad \ln W_j = \alpha + \beta_1 \ln\left(1 - q_j\right) + \beta_2 \ln d_j, \quad \beta_1 = 1, \quad j = 1,...,M,$$

where d_j denotes duration. An advantage of equation (6) is that it supports testing the method proposed here as well as testing a QALY model. If QALYs adequately represent preferences, then WTP is proportional to QALY losses and $\beta_1 = 1 = \beta_2$.Section 4C reports tests of the restrictions in equations (5) and (6).

Besides supporting a test for consistency with the underlying theory, incorporating information from multiple baselines would allow analysts to improve WTP estimation by, for example, constructing a range of estimates for sensitivity analysis, averaging baseline WTP estimates, or combining them to minimize the variance of the compound estimate. These possibilities, none of which is available for the valuation method developed by Kenkel (2006), are explored in the online Appendix.

[13] If $EU = \sum_{j=1}^{M} p_j U\left(I,S_j\right) + \left(1 - \sum_{j=1}^{M} p_j\right) U(I,H)$ then $W_j = \left[U(I,H) - U\left(I,S_j\right)\right] \big/ MEUI$, which together with equation (3) for all *j* implies equality of all $W_j \big/ \left(1 - q_j\right)$, where *MEUI* denotes the derivative of *EU* with respect to income.

For instance, suppose there are two baseline conditions with known preference weights and WTP (q_2, q_3, W_2, W_3) and one condition to be valued with known preference weight (q_1) but unknown WTP (W_1). The theoretical restriction that $W_2/(1-q_2) = W_3/(1-q_3)$ implies that either baseline could be used to compute the unknown WTP, yielding the same result. In practice, estimates of $W_j/(1-q_j)$ will differ across conditions. Then W_1 can be estimated as the following weighted average:

$$W_1 = (1-q_1)\left[\lambda \frac{W_2}{1-q_2} + (1-\lambda)\frac{W_3}{1-q_3}\right]. \quad (7)$$

Choosing λ to minimize the variance of the expression in brackets in equation (7) yields $\lambda = \sigma_{33}/(\sigma_{22}+\sigma_{33})$, where σ_{jj} denotes the variance of the estimator of $W_j/(1-q_j)$, $j = 2,3$.[14]

3. Methods

A. Estimating WTP

Our goal is to estimate WTP to avoid numerous morbidity conditions where it is currently unknown. As shown in equation (4), we require three inputs: WTP to avoid the baseline condition (W_2), the preference weight for the baseline condition, and preference weights for the morbidity conditions to be valued. This section describes our choices for these components, and discusses steps we take to assess potential measurement error from alternative methods to elicit preference weights.

[14] See the online Appendix. The result assumes that estimators of $W_2/(1-q_2)$ and $W_3/(1-q_3)$ are mutually uncorrelated (σ_{23}=0) and uncorrelated with the estimator of q_1. The approach generalizes to cases where $\sigma_{23} \neq 0$ or there are more than 2 baselines.

For the baseline condition, we select chronic kidney disease (CKD), using data from a comprehensive OECD study that evaluates CKD stages 3–5 combined, excluding dialysis (Dockins et al. 2024). WTP to avoid CKD was obtained using stated preference methods based on an online valuation survey (Dockins et al. 2024). Framing of the situation involved providing background information on the role of the kidneys, problems associated with kidney disease, and consequences of kidney failure. After this background, respondents were asked whether they would pay a specified cost to reduce risk of CKD by a specified amount. Costs and risk reductions were randomly assigned, with the reduction displayed graphically to improve communication. From this, the mean and median value of a statistical case of CKD was calculated. Although estimates are available for several countries, we limit our focus on mean estimates for the United States, and we show estimates using median CKD in the Appendix. While alternative baseline conditions could be considered, CKD is chosen due to the high quality and reliability of its measurement.

We obtain q_2, the preference weight for the baseline condition, from Gorodetskaya et al. (2005), who provide separate measurements for CKD stages 3, 4, and 5 using both Time Trade-Off (TTO) and Health Utilities Index (HUI) methods. We average across the 3 stages of CKD to obtain a measure comparable to WTP to avoid CKD.[15] Furthermore, to obtain a single preference weight, we average across the elicitation methods. This average can reduce errors by mitigating the impact of random or method-specific errors specific to each method, particularly if

[15] We compute the overall mean by weighting the mean in each stage by its sample size. We compute the overall standard deviation according to the following formula: $\sqrt{\left((n_1-1)s_1^2+(n_2-1)s_2^2+(n_3-1)s_3^2\right)/(n_1+n_2+n_3-3)}$, where n_k is the sample size and s_k is the standard deviation for stage k.

uncorrelated. For example, VAS may be influenced by scale anchoring, while TTO may reflect difficulty in understanding trade-offs. Averaging helps to balance out these biases, providing a more neutral estimate. We describe below how variations in these methods are used to assess potential measurement error in preference weights.

For q_1 the preference weights for the conditions we value, we utilize data from two key studies, Bell et al. (2001) and Salomon and Murray (2004), as both provide weights for multiple conditions using a consistent methodology.[16] Bell et al. (2001) conduct a systematic review of published preference scores. Salomon and Murray (2004) use a single sample and survey instrument to obtain preference weights for health conditions, thus providing consistency in survey interpretation by subjects, which limits the influence of individual-level biases or misunderstandings about the task. Using the same instruments also improves methodological comparability. Furthermore, Salomon and Murray (2004) employ multiple methods to collect these weights, where we use the same procedure for CKD by averaging the preference weights across measurement methods.

Combining these components, we compute the mean WTP to avoid the unknown conditions, with confidence intervals calculated using the delta method, shown in the online Appendix.

[16] Bell et al. (2001) report minimum and maximum values, from which we average to calculate the mean. Salomon and Murray (2004) report results in figures, from which we extract exact values using Plot Digitizer (https://plotdigitizer.com/app). Although somewhat dated, these studies were chosen because of the range of conditions with estimates available, which enables us to value multiple morbidity states. More recent preference weights can be chosen by the analyst for the particular condition of interest. We avoid using preference weights from the standardized EQ-5D instrument due to their limited variation and reduced ability to differentiate between health states. For example, when rank ordering the 95 health conditions evaluated by Sullivan and Ghushchyan (2006), the 5th percentile preference weight is 0.692 and 95th percentile is 0.826. Using equation (4) to value the 95 conditions with a given baseline condition, the 95th percentile WTP estimate would be only 119% of the 5th percentile estimate.

One potential limitation is that the preference weights for CKD are drawn from data collected in 2005 while the estimate for WTP to avoid CKD are 20 years later. To the extent that advances in treatment have altered the quality of life associated with CKD or other conditions, the resulting weights may not fully reflect current health states. This concern is particularly relevant if such changes differ across conditions, as our method relies on relative differences in preference weights. However, unless there have been substantial differential changes in quality of life across conditions, the resulting bias is likely to be limited.

B. Assessing Measurement Error from Preference Weights

Since preference weights are derived from stated preferences, they can be problematic for several reasons. Respondents may not fully understand the health conditions or trade-offs presented to them, leading to biased or inconsistent responses. Stated preferences are subject to framing effects, where the way questions are presented can significantly influence the outcomes. Hypothetical bias may arise whereby individuals respond differently in a survey setting compared to real-life situations involving actual trade-offs.[17] Consistent with this, Salomon & Murray show that different methods yield varying results, though these different methods generally preserve the rank order of conditions.

Treating this issue as one of measurement error, we assess the sensitivity of our results to different measures of the preference weights. We calculate two separate WTP estimates for the unknown conditions by using the same methods for eliciting preference weights rather than

[17] For example, Salomon and Murray note that maintaining a consistent duration for all health states in their study may create some unrealistic hypothetical scenarios, such as two broken arms in stiff casts for 10 years.

averaging across them. Specifically, we apply Equation (4) using only Time Trade-Off (TTO) estimates for q_1 and q_2, and repeat the calculation using Visual Analog Scale (VAS) estimates for q_1 and q_2.[18] This yields a range of estimates for each condition that we can use to quantitatively assess the degree of measurement error induced by using stated preference measures for the preference weights.[19]

C. Assessing Validity

We use data from the meta-analysis of Johnson, Fries and Banzhaf (1997; hereafter, JFB) to assess the validity of our method. Table 1 in JFB reports preference weights and WTP estimates (in 1993$) for acute conditions of varying duration (1, 2, 7, 10, 20, 30, and 90 days). The conditions range from minor ailments like throat congestion or runny nose to more serious problems like severe shortness of breath or severe angina symptoms. The data include a total of 53 condition-duration combinations , where each observation consists of a willingness-to-pay (WTP) estimate (in 1993 dollars), a corresponding preference weight based on the Quality of Well Being (QWB), and the duration of the condition. WTP estimates (in 1993 dollars) were taken from five stated-preference studies conducted in the 1970s and 1980s; because the analysis uses logarithms, results are invariant to the choice of base year.

[18] Note that we use HUI-3 for q_2, which we contend is closest to VAS. The HUI procedure involves respondents rating their health across eight attributes, with multiple states provided for each attribute. Preference weights for each state were derived using VAS, with some states also using SG to limit some concerns associated with VAS (Horsman et al. 2003).

[19] Although WTP to avoid CKD also is elicited by stated preference, we focus on measurement error in q because multiple methods are used often to elicit weights (Bell et al. 2001), but rarely to elicit WTP.

We use the JFB data to test our model's prediction that the elasticity of WTP with respect to (*1-QWB*) is unity, and the additional restriction of a QALY model that the elasticity of WTP with respect to duration is unity. Specifically, we specify equations (5) and (6) as linear-in-log regressions, with log WTP as the dependent variable and log(1−QWB) and log(duration) as key independent variables (using these 53 observations), along with indicator variables to control for the five original studies that produced the WTP estimates. We test whether the estimated coefficients on log(1−QWB) and log(duration) are equal to one, as implied by the theoretical restrictions.

4. Results

A. Main Results

Table 1 presents estimated preference weights and WTP along with 95% confidence intervals. For the baseline condition CKD, the average preference weight from Gorodetskaya et al. (2005) is q_2 = 0.72 (CI 0.68, 0.77). Preference weights for the conditions to be valued using equation (4) range from 0.29 (quadriplegia) to 0.9 (chronic hepatitis).[20] More severe conditions have lower weights, consistent with expectations. For example, kidney dialysis (CKD stage 5) has a lower weight than CKD stages 3-5 (0.55 < 0.72).

Mean WTP to avoid the baseline condition of CKD from Dockins et al. (2024) is W_2 = \$0.68 million (CI \$0.64m-\$0.72m). Applying equation (4), we generate WTP estimates ranging from approximately \$0.26 million (chronic hepatitis) to \$1.74 million (quadriplegia). Estimates

[20] We value morbidity for all conditions considered by Bell et al. (2001) and Salomon and Murray (2004) to illustrate application of our method across a diverse set of conditions.

based on median WTP to avoid CKD are reported in Appendix Table A1 and are 22% of mean-based estimates in Table 1. Confidence intervals for estimated WTP are reasonably narrow. The ratio of width of interval to WTP ranges from 0.38 to 1.84, with a median of 0.63.

Ten of 18 WTP estimates exceed $1 million but are relatively modest compared to VSLs used by federal agencies of approximately $12 million (Federal Register 2024) and to prior estimates of WTP to avoid morbidity. For instance, our WTP to avoid chronic bronchitis ($0.56 million) is 52% of the estimate of Viscusi, Magat and Huber (1991; inflated to 2022$ using CPI-U).[21] In a study of WTP per QALY, Hammitt and Haninger (2017) estimated the value of avoiding a health state lasting 33 years with $(1\text{-}q) = 0.26$ at $1.92 million in 2022$. Our estimates for conditions with similar values of $1\text{-}q$ (0.23 to 0.29) range from $0.56 to $0.72m; the largest of the upper confidence limits is only 38% of the Hammitt and Haninger estimate.

In contrast, our WTP estimates substantially exceed illness costs estimates used to value air pollution regulations (EPA 2024). For mild angina, mean WTP is about $0.40 million compared to $48,000 illness costs; for major stroke, WTP is $1.6 million compared to less than $35,000.[22] These disparities suggest substantial undervaluation of morbidity in current regulatory analyses.

We also provide WTP estimates for illnesses related to conditions of interest noted by the recent report of the Advancing Frontiers of BCA, including liver disease and mental health

[21] The median value from Viscusi, Magat and Huber, updated to 2022$, is $1.08m. Our median based value (Appendix Table A1) is $0.12 million.

[22] Our median estimates in Table A1 are $95,000 ($25,000-$0.16m) for angina and $0.35 million ($0.25m-$0.45m) for major stroke.

outcomes. We estimate WTP values of $1.7 million for avoiding liver cancer, $1.69 million for avoiding active psychosis and $1.57 million for avoiding major depression.

B. Measurement Error in Preference Weights

Table 2 presents mean WTP estimates and confidence intervals using alternative preference weights; comparable results for median WTP are shown in Appendix Table A2. Using TTO alone to measure weights, rather than averaging over methods as in Table 1, WTP estimates are 106% to 182% of Table 1 values. Alternatively, if the CKD weight is based on HUI-3 and weights for other conditions are based on VAS, WTP estimates are 73% to 84% of values in Table 1.

WTP computed from TTO ranges from 1.28 to 2.43 times the WTP computed from HUI-3 and VAS, with 7 of 12 conditions having TTO-based WTP less than double HUI-3/VAS-based WTP. We compute an upper bound on sensitivity of WTP estimates to alternative preference weights as the ratio of the upper confidence limit of TTO-based WTP to the lower limit of HUI-3/VAS-based WTP (in effect considering a probability of 0.025^2=0.000625), which ranges from 2.7 to 3.5. These results suggest that the sensitivity of WTP estimates to different methods of measuring preference weights may often be less than a factor of 2, and rarely more than a factor of 3.5. In comparison, Viscusi and Aldy (2003) report VSL estimates that range by a factor of at least 2.25.

C. Testing the Theory

Using JFB data, we estimate equation (6) with indicator variables to control for the five original studies that produced the WTP estimates (95% CI in parentheses):

$$\widehat{\ln(WTP)} = \underset{(0.617,\ 1.951)}{1.284} \times \ln(1-QWB) + \underset{(0.409,\ 0.599),}{0.504} \times \ln(Days) + \text{study indicator variables} \qquad F(7,46) = 593.458.$$

As expected, WTP increases with the relative disutility and duration. The null hypothesis that the coefficient of *ln(1-QWB*) equals one is not rejected (p=0.41), consistent with our model. But a unit elasticity for duration, necessary for consistency with QALYs as measures of preferences, is rejected at less than one percent significance. Appendix Table A3 reports results from two additional specifications, one with no control for duration (like equation (5)) and one with dummy variables for duration. The hypothesis of a unit elasticity of WTP with respect to (*1-QWB*) is not rejected in either case (p > 0.6), providing further support for the theoretical prediction that $W_j \big/ \left(1-q_j\right)$ is constant over conditions.[23]

[23] Other researchers have performed similar tests to investigate the validity of QALYs as preference measures, but none have tested proportionality of WTP to $(1-q)$ as the key theoretical restriction supporting morbidity valuation using equation (4). In their meta-analysis, JFB (1992) found that WTP increased less than proportionately with QWB (not $1-QWB$). Pinto-Prades et al. (2009) used individual data and found that WTP was less than proportional to $(1-q)$ based on EQ-5D. Haninger and Hammitt (2011, using HUI-3) and Hammitt and Haninger (2017, using EQ-5D) also used individual data and found that WTP increases less than proportionately with $\left(q_0 - q\right)$, where q_0 is the preference weight corresponding to an individual's current health. Each of these studies also found that WTP is less than proportional to duration. Van Houtven et al. (2006) used QWB in their meta-analysis and found that WTP is less than proportional to duration but increases more than proportionately with $(1-q)$.

5. Conclusion

This study introduces a theoretically consistent and practical framework for valuing morbidity that overcomes the limitations of existing approaches such as cost-of-illness (COI) measures and Quality-Adjusted Life Years (QALYs). Building on an expected utility model, the method derives willingness-to-pay (WTP) estimates for specific health conditions based on preference weights and a baseline condition with a known WTP. Rather than attempting to produce an exhaustive list of morbidity valuations, we provide a handful of illustrative estimates under various assumptions, using chronic kidney disease (CKD) as the baseline condition. The results highlight significantly higher WTP estimates for health states, such as mild angina and major stroke, compared to COI measures commonly used in policy evaluations. We assess the robustness of this method through sensitivity analyses, accounting for variations in utility weights and baseline conditions, and obtain variation in WTP estimates comparable to those observed for the Value of Statistical Life (VSL), widely used in policy analysis. Additional results support the consistency of the method with theoretical restrictions, suggesting that it provides a reasonable substitute for direct estimation of WTP when primary research is not feasible.

This method is straightforward to apply, offering practical guidelines for valuing morbidity in various policy contexts. First, identify the preference weights for the health condition of interest; the Tufts Cost-Effectiveness Analysis Registry (https://cear.tuftsmedicalcenter.org/) provides a valuable starting point for locating these weights. Next, combine the identified preference weight with the willingness-to-pay (WTP) estimate for chronic kidney disease (CKD) provided in this paper, or select another baseline condition with a known WTP. Since WTP varies with duration but preference weights do not,

choose a chronic or acute baseline according to whether the condition to be valued is chronic or acute. When choosing preference weights, use those derived from the same elicitation method (e.g., Visual Analog Scale or Time Trade-Off) as the baseline condition. If multiple preference weights are available, prioritize those that most closely match the health state characteristics of the baseline condition, then average across them. Alternatively, assess the sensitivity of the estimates by using different preference weights for the condition being valued. These practical steps make the method accessible and flexible for a range of applications, enabling policymakers and researchers to generate robust morbidity valuations efficiently.

We readily acknowledge areas for refinement but stress the immediate utility of the approach in improving the accuracy and comprehensiveness of health-related policy evaluations. For example, our approach assumes the independent valuation of health outcomes, consistent with how preference weights are derived. Consequently, this limits the applicability of our method in scenarios involving co-morbidities or the simultaneous occurrence of multiple conditions. Further efforts to enhance this method and to assess its utility in various settings involving prevention, treatment, and possibly damage assessment, alongside the development of analogously defined preference weights, represent a promising avenue for future research.

REFERENCES

Bell, Chaim M, Richard H Chapman, Patricia W Stone, Eileen A Sandberg, and Peter J Neumann. 2001. "An Off-the-Shelf Help List: A Comprehensive Catalog of Preference Scores from Published Cost-Utility Analyses." *Medical Decision Making* 21: 288-294.

Bleichrodt, Han, and John Quiggin. 1999. "Life-Cycle Preferences over Consumption and Health: When Is Cost-Effectiveness Analysis Equivalent to Cost–Benefit Analysis?" *Journal of Health Economics* 18 (6): 681–708. https://doi.org/10.1016/S0167-6296(99)00014-4.

Cropper, Maureen, Emily Joiner and Alan Krupnick. 2024. "Revisiting EPA's Value per Statistical Life." *Review of Environmental Economics and Policy* 18(2): 193-211.

Dockins, Chris, Damien Dussaux, Charles Griffiths, Nathalie Simon and Sandra Hoffmann. 2023. "Valuing a Reduction in the Risk of Chronic Kidney Disease: A Large-Scale Multi-Country Stated-Preference Approach." OECD Environmental Working Papers No. 216. https://www.oecd.org/en/publications/valuing-a-reduction-in-the-risk-of-chronic-kidney-disease_9c93138f-en.html.

Dockins, Chris, Damien Dussaux, Charles Griffiths, Sandra Hoffmann, and Nathalie Simon. 2024. "Valuing a Reduction in the Risk of Chronic Kidney Disease: A Large-Scale Multi-Country Stated Preference Approach." *Journal of Benefit-Cost Analysis*, September, 1–21. https://doi.org/10.1017/bca.2024.16.

Federal Register. 2024. "Notice of Availability of Final Guidance for Estimating Value per Statistical Life." Consumer Product Safety Commission. Vol. 89, No. 76 (April 18): 27740-27751.

Gorodetskaya, Irina, Stefanos Zenios, Charles E. Mcculloch, Alan Bostrom, Chi-Yuan Hsu, Andrew B. Bindman, Alan S. Go, and Glenn M. Chertow. 2005. "Health-Related Quality of Life and Estimates of Utility in Chronic Kidney Disease." *Kidney International* 68 (6): 2801–8. https://doi.org/10.1111/j.1523-1755.2005.00752.x.

Hammitt, James K. 2013. "Admissible Utility Functions for Health, Longevity, and Wealth: Integrating Monetary and Life-Year Measures." *Journal of Risk and Uncertainty* 47 (3): 311–25.

Hammitt, James K. 2017. "Valuing Non-fatal Health Risks: Monetary and Health-Utility Measures." *Revue économique* 68(3): 335-356.

Hammitt, James K., and Kevin Haninger. 2017. "Valuing Nonfatal Health Risk as a Function of Illness Severity and Duration: Benefit Transfer Using QALYs." *Journal of Environmental Economics and Management* 82 (March):17–38. https://doi.org/10.1016/j.jeem.2016.10.002.

Haninger, K., Hammitt, J.K., 2011. Diminishing willingness to pay per quality-adjusted life year: valuing acute foodborne illness. Risk Anal. 31, 1363–1380.

Harrington, W., Portney, P., 1987. Valuing the benefits of health and safety regulation. Journal of Urban Economics 22(1), 101-112.

Horsman, J., Furlong, W., Feeny, D. et al. 2003. "The Health Utilities Index (HUI): concepts, measurement properties and applications." *Health Qual Life Outcomes* 1, 54. https://doi.org/10.1186/1477-7525-1-54.

Institute of Medicine. 2006. *Valuing Health for Regulatory Cost-Effectiveness Analysis*. Washington, D.C.: National Academies Press. https://doi.org/10.17226/11534.

Johnson, Reed F., Erin E. Fries, and H. Spencer Banzhaf. 1997. "Valuing Morbidity: An Integration of the Willingness-to-Pay and Health-Status Index Literatures." *Journal of Health Economics* 16 (6): 641–65. https://doi.org/10.1016/S0167-6296(97)00012-X.

Johnston, Robert J., Kevin J. Boyle, Maria L. Loureiro, Ståle Navrud, and John Rolfe. 2021. "Guidance to Enhance the Validity and Credibility of Environmental Benefit Transfers." *Environmental and Resource Economics* 79 (3): 575–624. https://doi.org/10.1007/s10640-021-00574-w.

Kenkel, Don. 2006. "WTP- and QALY-Based Approaches to Valuing Health for Policy: Common Ground and Disputed Territory." *Environmental & Resource Economics* 34 (3): 419–37. https://doi.org/10.1007/s10640-006-9003-5.

Miyamoto, John M., Peter P. Wakker, Han Bleichrodt, and Hans J. M. Peters. 1998. "The Zero-Condition: A Simplifying Assumption in QALY Measurement and Multiattribute Utility." *Management Science* 44 (6): 839–49. https://doi.org/10.1287/mnsc.44.6.839.

Pinto-Prades, J.L., Loomes, G., Brey, R., 2009. Trying to estimate a monetary value for the QALY. J. Health Econ. 28, 553–562.

Rosenberger, Randall.S., Loomis, John.B. 2017. "Benefit Transfer." In: Champ, P., Boyle, K., Brown, T. (eds) *A Primer on Nonmarket Valuation*. The Economics of Non-Market Goods and Resources, vol 13. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7104-8_11.

Salomon, Joshua A., and Christopher J.L. Murray. 2004. "A Multi-method Approach to Measuring Health-state Valuations." *Health Economics* 13 (3): 281–90. https://doi.org/10.1002/hec.834.

Sullivan, Patrick and and Vahram Ghushchyan. 2006. "Preference-Based EQ-5D Index Scores for Chronic Conditions in the United States." *Medical Decision Making* 26: 410-420.

Torrance, George W. 1986. "Measurement of Health State Utilities for Economic Appraisal: A Review." *Journal of Health Economics* 5: 1-30.

US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. 2017. *Guidelines for Regulatory Impact Analysis*. *https://aspe.hhs.gov/reports/guidelines-regulatory-impact-analysis*

US Environmental Protection Agency, Office of Air and Radiation. 2024. *Estimating $PM_{2.5}$ and Ozone-Attributable Health Benefits: 2024 Update*. Research Triangle Park, North Carolina. June. https://www.epa.gov/system/files/documents/2024-06/estimating-pm2.5-and-ozone-attributable-health-benefits-tsd-2024.pdf.

US National Science and Technology Council, Subcommittee on Frontiers of Benefit-Cost Analysis. 2023. *Advancing the Frontiers of Benefit-Cost Analysis: Federal Priorities and Directions for Future Research*. December. https://www.whitehouse.gov/wp-content/uploads/2023/12/FINAL-SFBCA-Annual-Report-2023.pdf.

US Office of Management and Budget. 2023. Circular No. A-4. Nov. 9. https://www.whitehouse.gov/wp-content/uploads/2023/11/CircularA-4.pdf.

Van Houtven, G., Powers, J., Jessup, A., Yang, J.-C., 2006. Valuing avoided morbidity using meta-regression analysis: what can health status measures and QALYs tell us about willingness-to-pay? Health Econ. 15, 775–795.

Viscusi, W Kip, Wesley A Magat, and Joel Huber. 1991. "Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-Offs for Chronic Bronchitis." *Journal of Environmental Economics and Management* 21 (1): 32–51. https://doi.org/10.1016/0095-0696(91)90003-2.

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

Table 1. Morbidity Values based on Mean WTP to Reduce Risk of Chronic Kidney Disease.

	Preference Weights			Willingness to Pay ($1,000)		
		95% Confidence Interval			95% Confidence Interval	
Condition	Mean	Lower	Upper	Mean	Lower	Upper
CKD	0.72	0.68	0.77	$ 677	$ 637	$ 717
Mild angina	0.83	0.70	0.95	$ 429	$ 114	$ 743
Liver cancer	0.30	0.10	0.49	$ 1,727	$ 1,167	$ 2,286
Major stroke	0.35	0.20	0.50	$ 1,592	$ 1,137	$ 2,047
Chronic hepatitis	0.90	0.80	0.99	$ 257	$ 21	$ 494
Kidney dialysis	0.55	0.41	0.68	$ 1,114	$ 734	$ 1,495
Below knee amputation	0.71	0.61	0.80	$ 723	$ 460	$ 985
Quadriplegia	0.29	0.23	0.37	$ 1,739	$ 1,411	$ 2,067
Active psychosis	0.31	0.24	0.39	$ 1,690	$ 1,357	$ 2,023
Major depression	0.36	0.28	0.44	$ 1,568	$ 1,238	$ 1,897
Hemiparesis	0.43	0.35	0.52	$ 1,396	$ 1,090	$ 1,703
Below knee amputation, both legs	0.45	0.36	0.54	$ 1,347	$ 1,030	$ 1,664
Two broken arms	0.49	0.41	0.59	$ 1,249	$ 961	$ 1,537
Blindness	0.50	0.40	0.59	$ 1,225	$ 904	$ 1,545
Below knee amputation, one leg	0.60	0.50	0.68	$ 980	$ 684	$ 1,275
Deafness	0.71	0.62	0.79	$ 710	$ 459	$ 961
Chronic bronchitis	0.77	0.68	0.84	$ 563	$ 323	$ 803
Watery diarrhoea	0.84	0.76	0.89	$ 392	$ 185	$ 599
Vitiligo on face	0.89	0.83	0.94	$ 269	$ 116	$ 423

Notes: All results reported in 2022$. Sources for preference weights are Gorodetskaya et al. (2005) for CKD, Bell et al. (2001) for mild angina through below knee amputation, Salomon and Murray (2004) for quadriplegia through vitiligo on face. WTP to avoid CKD is from Dockins et al. (2024). Remaining WTP estimates were computed by authors applying equation (4), with standard errors estimated by the delta method.

Table 2. Mean WTP and 95% Confidence Interval based on Alternative Measures of Preference Weights.

	Estimates based on TTO ($1,000)			Estimates based on VAS and HUI-3 ($1,000)		
	Mean	Lower	Upper	Mean	Lower	Upper
Quadriplegia	$ 3,080	$ 1,821	$ 4,338	$ 1,345	$ 1,067	$ 1,622
Active psychosis	$ 3,073	$ 1,821	$ 4,325	$ 1,266	$ 990	$ 1,542
Major depression	$ 2,640	$ 1,404	$ 3,875	$ 1,164	$ 873	$ 1,455
Hemiparesis	$ 2,129	$ 1,023	$ 3,234	$ 1,033	$ 762	$ 1,304
Below knee amputation, both legs	$ 2,012	$ 1,016	$ 3,008	$ 1,033	$ 812	$ 1,254
Two broken arms	$ 1,850	$ 954	$ 2,746	$ 968	$ 738	$ 1,197
Blindness	$ 1,857	$ 747	$ 2,966	$ 902	$ 624	$ 1,180
Below knee amputation, one leg	$ 1,281	$ 716	$ 1,846	$ 778	$ 539	$ 1,017
Deafness	$ 938	$ 571	$ 1,305	$ 569	$ 299	$ 839
Chronic bronchitis	$ 776	$ 504	$ 1,049	$ 473	$ 302	$ 644
Watery diarrhoea	$ 421	$ 262	$ 579	$ 329	$ 170	$ 488
Vitiligo on face	$ 285	$ 166	$ 403	$ 216	$ 156	$ 275

Notes: All results reported in 2022$. For baseline condition (CKD), reference weights are from Gorodetskaya et al. (2005) and mean WTP is from Dockins et al. (2024). For other conditions, preference weights are from Salomon and Murray (2004) and WTP estimates were computed by authors applying equation (4), with standard errors estimated by the delta method.

Appendix: Supplemental Results

Table A1. Morbidity Values based on Median WTP to Reduce Risk of Chronic Kidney Disease.

	Preference Weights			Willingness to Pay ($1,000)		
		95% Confidence Interval			95% Confidence Interval	
Condition	Mean	Lower	Upper	Median	Lower	Upper
CKD	0.72	0.68	0.77	149	149	150
Mild angina	0.83	0.70	0.95	95	25	164
Liver cancer	0.30	0.10	0.49	381	258	505
Major stroke	0.35	0.20	0.50	351	251	452
Chronic hepatitis	0.90	0.80	0.99	57	5	109
Kidney dialysis	0.55	0.41	0.68	246	162	330
Below knee amputation	0.71	0.61	0.80	159	101	217
Quadriplegia	0.29	0.23	0.37	384	311	456
Active psychosis	0.31	0.24	0.39	373	300	446
Major depression	0.36	0.28	0.44	346	273	419
Hemiparesis	0.43	0.35	0.52	308	240	376
Below knee amputation, both legs	0.45	0.36	0.54	297	227	367
Two broken arms	0.49	0.41	0.59	276	212	339
Blindness	0.50	0.40	0.59	270	200	341
Below knee amputation, one leg	0.60	0.50	0.68	216	151	281
Deafness	0.71	0.62	0.79	157	101	212
Chronic bronchitis	0.77	0.68	0.84	124	71	177
Watery diarrhoea	0.84	0.76	0.89	86	41	132
Vitiligo on face	0.89	0.83	0.94	59	26	93

Notes: This table presents WTP estimates based on applying equation (4) to median WTP for CKD, for comparison to the estimates based on mean WTP for CKD that were presented in Table 1 in the paper. All results reported in 2022$. Sources for preference weights are Gorodetskaya et al. (2005) for CKD, Bell et al. (2001) for mild angina through below knee amputation, Salomon and Murray (2004) for quadriplegia through vitiligo on face. Median WTP for CKD is from Dockins et al. (2024). Remaining WTP estimates were computed by authors applying equation (4), with standard errors estimated by the delta method.

Table A2. Median WTP and 95% Confidence Interval based on Alternative Measures of Preference Weights.

	Estimates based on TTO ($1,000)			Estimates based on VAS and HUI-3 ($1,000)		
	Mean	Lower	Upper	Mean	Lower	Upper
Quadriplegia	678	401	954	296	235	357
Active psychosis	676	401	952	279	218	339
Major depression	581	309	853	256	192	320
Hemiparesis	468	225	712	227	168	287
Below knee amputation, both legs	443	224	662	227	179	276
Two broken arms	407	210	604	213	162	263
Blindness	409	164	653	198	137	260
Below knee amputation, one leg	282	158	406	171	118	224
Deafness	206	126	287	125	66	185
Chronic bronchitis	171	111	231	104	66	142
Watery diarrhoea	93	58	127	72	37	107
Vitiligo on face	63	36	89	47	34	61

Notes: This table illustrates the sensitivity of the median-based WTP estimates to different methods of measuring preference weights. Table 2 in the paper presented comparable results for the mean-based WTP estimates. All results reported in 2022$. For baseline condition (CKD), reference weights are from Gorodetskaya et al. (2005) and median WTP is from Dockins et al. (2004). For other conditions, preference weights are from Salomon and Murray (2004) and WTP estimates were computed by authors applying equation (4), with standard errors estimated by the delta method.

Table A3: Testing the Model with Data from JFB.

	Dependent variable: ln(WTP)		
	(1)	(2)	(3)
ln(1 - QWB)	1.284	1.303	1.132
	(0.340)	(0.617)	(0.354)
ln(Days)	0.504		
	(0.048)		
Duration indicators:			
1 day			-1.943
			(0.275)
2 days			-1.443
			(0.765)
7 days			-1.017
			(0.289)
10 days			-0.667
			(0.494)
20 days			0.492
			(0.494)
30 days			-0.060
			(0.385)
90 days (excluded category)			---

Study-author effects:			
Chestnut et al. (1988)	6.266	6.460	7.978
	(0.340)	(0.911)	(0.683)
Dickie et al. (1987, 1988)	4.613	4.640	6.338
	(0.511)	(0.927)	(0.603)
Loehman et al. (1979)	4.904	6.011	6.785
	(0.451)	(0.795)	(0.484)
Rowe & Chestnut (1985)	6.700	7.857	8.328
	(0.653)	(1.167)	(0.855)
Tolley et al. (1986)	5.844	6.736	7.434
	(0.450)	(0.802)	(0.563)

Notes: This table presents estimation results for three specifications of the regression analysis used to test the theoretical restriction that the elasticity of WTP with respect to (1 – QWB) is unity. The first column of results is the specification reported in the paper. Standard errors in parentheses. 53 observations on: WTP to avoid various acute conditions in 1993$; QWB=Quality of Well Being preference weight; Days=duration of condition. Data source: Johnson, Fries and Banzhaf (1997), Table 1.