Willingness to Pay for Quality-Adjusted Life Years: Empirical Inconsistency Between Cost-Effectiveness Analysis and Economic Welfare Theory

Kevin Haninger
Center for Risk Analysis, Harvard University
haninger@fas.harvard.edu

James K. Hammitt
Center for Risk Analysis, Harvard University
IDEI and LERNA-INRA, Université de Toulouse
jkh@harvard.edu

February 2006

Abstract

We design and conduct a stated-preference survey to test whether willingness to pay (WTP) to reduce risk of acute illness is proportional to the corresponding change in expected quality-adjusted life years (QALYs). For the short-term illnesses we consider, proportionality is required by economic theory if QALYs measure utility for health. Proportionality implies a constant WTP per QALY and that WTP is proportional to changes in both health quality and duration of illness. WTP is elicited using double-bounded, dichotomous-choice questions in which respondents (randomly selected from the United States general adult population, n = 2,795) decide whether to purchase a more expensive food to reduce the risk of foodborne illness. Health risks vary by baseline probability of illness, reduction in probability, duration and severity of illness, and conditional probability of mortality. The expected gain in QALYs is calculated using respondent-assessed decrements in health-related quality of life if ill combined with the stated duration of illness and reduction in probability. We reject the hypothesis that WTP is proportional to changes in expected QALYs and find diminishing marginal WTP for severity and duration of illness prevented. Our results suggest that individuals do not have a constant rate of WTP for changes in QALYs, which implies that cost-effectiveness analysis using cost per expected QALY gained is not consistent with economic welfare theory.

Keywords: Quality-adjusted life year, willingness to pay, health risk, stated-preference, cost-effectiveness analysis, foodborne illness

1. Introduction

Willingness to pay (WTP) and quality-adjusted life years (QALYs) represent alternative methods of valuing health risk, but there is little empirical research on the relationship between them. How they relate has important implications for the role of cost-effectiveness analysis in evaluating policies that affect health and safety. Cost-effectiveness analysis, which evaluates interventions in terms of the cost per expected QALY gained, would be consistent with economic welfare theory and benefit-cost analysis if QALYs were a valid measure of utility and the WTP per QALY gained were constant in the population (Johannesson, 1995; Garber and Phelps, 1997; Hammitt, 2002a, 2002b). While it is clear that individual WTP per QALY is likely to vary with income and wealth, we consider the more fundamental question of whether an individual's WTP for risk reduction is proportional to the expected gain in QALYs. Our results suggest it is not, and hence cost-effectiveness analysis is not consistent with economic welfare theory.

We design and conduct a stated-preference survey to test whether WTP for a reduction in the risk of acute illness is proportional to the change in expected health and duration of illness, and hence proportional to the change in expected QALYs. The survey was completed by 2,795 randomly-selected adults in the United States. We reject the hypothesis that WTP is proportional to the loss of health-related quality of life (HRQL) and duration of illness and find that respondents exhibit diminishing marginal WTP for changes in these attributes.

In the following section, we describe the theoretical and empirical background for the study. First, we briefly review the literature on the value of reducing morbidity risk, focusing on studies that are relevant to QALY measures of health. Second, we describe the implications of economic theory for WTP per QALY and develop an empirical model to test for proportionality. In Section 3, we describe the survey instrument and sample. In Section 4, we report the results of regression models relating WTP to the severity and duration of illness, reduction in its probability, other risk attributes, and to demographic and preference characteristics of the respondents. We conclude in Section 5.

2. Background

In this section, we describe prior estimates of WTP per QALY, the implications of economic theory for the relationship between WTP and expected QALYs gained, and the empirical model we use to investigate this relationship.

2.1. Prior Work

Preliminary efforts to explore the relationship between WTP and QALYs have emerged within the last several years (O'Brien and Viramontes, 1994; Krabbe et al., 1997; Bala et al., 1998). Tolley et al. (1994) place dollar values on alleviation of a wide range of health conditions by assuming that WTP is proportional to QALYs gained. For example, they estimate WTP to alleviate a day of earache as the loss in HRQL (0.16) times the value of a statistical life-year ($120,000) divided by the number of days in a year, or $55. In direct elicitations, they observe that WTP is less than proportional to duration of improved health. It is unclear if this finding reflects insensitivity of the contingent valuation (CV) method to the scope of the good being valued or genuine preferences.

Jones-Lee et al. (1995) elicit WTP to reduce the probabilities of traffic fatality and of six specified traffic injuries of varying severity, and elicit probabilities of fatality for a standard gamble between fatality and complete recovery that respondents view as indifferent to each specified injury. They find that the ratio of WTP to prevent injury to WTP to prevent fatality is much larger than the standard gamble probability for the corresponding injury. They suggest this finding represents insensitivity of the WTP values to injury severity, although nonlinearity between WTP and HRQL could also contribute.

In a meta-analysis of five health-related CV studies, Johnson et al. (1997) examine WTP to avoid morbidity associated with 53 short-term health conditions. The authors assign an estimate of HRQL to each health condition using the Quality of Well Being Scale (QWB) (Kaplan et al., 1993), and estimate a generalized value function that places a dollar value on avoiding any short-term health condition that can be assigned a QWB score. The results show that WTP increases with both severity and duration of the illness avoided, although WTP is relatively insensitive to health quality for mild illnesses. Johnson et al. find evidence of increasing marginal WTP with the severity and decreasing marginal WTP with the duration of illness avoided. However, because rating-scale estimates of HRQL (such as the QWB) tend to exhibit an increasing convex relationship to standard-gamble and time-tradeoff estimates, this result does not preclude a linear relationship between WTP and other measures of HRQL.

Van Houtven et al. (2003) expand the Johnson et al. (1997) meta-analysis. They also find that WTP increases more than proportionally with severity and less than proportionally with the duration of illness averted. In addition, the authors find that avoiding loss of mobility and physical activity has significant and positive effects on WTP, while avoidance of symptoms and limitations in social activity have smaller, insignificant effects on WTP.

Johnson et al. (2000) apply discrete-choice modeling to estimate WTP to reduce the risk of suffering acute episodes of respiratory and cardiovascular illness. They find that WTP increases less than proportionally with the duration of illness averted. Gyrd-Hansen (2003) uses a similar technique to estimate WTP for improvements in chronic health in a random sample of the Danish population, using the EQ-5D (EuroQol Group, 1990) rather than the QWB to measure health. Both studies find evidence of variation in WTP for different attributes of health. In addition, Gyrd-Hansen's results suggest that WTP to improve health is not proportional to the EQ-5D scoring function.

In summary, previous studies have found that marginal WTP to prevent an illness decreases with the duration of the illness avoided (Tolley et al., 1994; Johnson et al., 1997; Johnson et al., 2000; Van Houtven et al., 2003) but increases with severity (Johnson et al., 1997; Van Houtven et al., 2003).

2.2. Theory and Empirical Model

WTP and QALYs represent alternative methods of valuing reductions in health risk. WTP is defined as the compensating variation, i.e., the maximum amount of money an individual would exchange for the reduction in health risk. A closely related measure, willingness to accept

(WTA), is defined as the equivalent variation, i.e., the minimum amount of money an individual would accept to forgo the reduction in health risk. For small changes in risk, WTP and WTA should be nearly equal in magnitude, although differences can arise for large changes in risk or when the risk change has no close substitutes (Hanemann, 1991).

QALYs are an alternative measure of individual utility for health that imposes additional structure on preferences. If an individual's preferences for health are consistent with QALYs, they must satisfy several conditions, including mutual utility independence of health and longevity, constant proportional tradeoff of longevity for health, and risk neutrality for longevity (Pliskin et al., 1980), or risk neutrality with respect to longevity for any health state and indifference to health quality for periods of zero duration (Bleichrodt et al., 1997). In addition, preferences for tradeoffs between health and longevity are assumed to be independent of income (Hammitt, 2002a).

Standard economic theory places few constraints on how WTP varies with health risk, except that WTP should increase with the magnitude of the benefit and, for small reductions in the probability of harm, WTP should be nearly proportional to the change in probability (Hammitt, 2000; Corso et al., 2001).

Consider an individual facing a probability r_0 of suffering an acute illness. Her current HRQL is h_0; if she becomes ill, it will fall to $h_1 < h_0$ for the duration of the illness and then return to h_0. She is offered the opportunity to reduce the chance of falling ill to $r_1 < r_0$. If her WTP to reduce the risk of illness is proportional to the gain in expected QALYs, then her WTP v for this risk reduction is given by

$$v = r\ h\ t\ m \tag{1}$$

where $r = r_0 - r_1$ is the reduction in probability of illness, $h = h_0 - h_1$ is the decrement in HRQL while ill, t is the duration of the illness, and m is her rate of substitution between QALYs and income. We test for proportionality of WTP to the expected change in QALYs, and to each of the factors comprising it, by embedding Equation (1) in a more general regression model that allows for non-proportionality of WTP to the risk reduction, decrement in HRQL, and duration of illness. We estimate the following generalization of the logarithm of Equation (1),

$$\log(v) = \delta \log(r) + \alpha \log(h) + \beta \log(t) + \gamma \log(m_0) + \sum_{j=1}^{n} \theta_j X_j + \varepsilon \tag{2}$$

where X is an n-dimensional vector of characteristics of the respondent and health risk that may influence m, m_0 is the rate of substitution between QALYs and income when $X_j = 0$ for all j, and ε is a mean-zero error term.

Proportionality of WTP to the change in expected QALYs requires $\alpha = \beta = \delta = 1$. If α is significantly different from one, we reject the hypothesis of proportionality of WTP to HRQL. A value of $\alpha < 1$ indicates diminishing marginal WTP to prevent an HRQL decrement (increasing marginal WTP for health improvement) and $\alpha > 1$ indicates increasing marginal WTP for the averted health decrement (decreasing marginal WTP for health improvement).

Similarly, if β is significantly different from one, we reject the assumption of a linear relationship between WTP and duration of illness. If $\beta < 1$, WTP is less than proportional to duration of illness averted, and $\beta > 1$ implies that WTP increases more than proportionally with duration. The QALY formulation evaluates changes in health as the product ht, and hence requires $\alpha = \beta = 1$.

If WTP is proportional to the change in probability of illness, $\delta = 1$. For small changes in risk, near proportionality of WTP to probability of harm is implied by expected utility theory and most alternative theories of decision making under uncertainty (e.g., cumulative prospect theory; Kahneman and Tversky, 1992) except when there is a discontinuity in the probability-weighting function for the relevant probabilities (see Hammitt, 2000 or Corso et al., 2001 for discussion). For large changes in risk, WTP should be less than proportional to risk reduction because of the income effect (which reduces marginal WTP for successive risk increments). Marginal WTP could also decrease for large risk reductions because of an analog of the "dead-anyway" effect (Pratt and Zeckhauser, 1996) if the marginal utility of income is smaller when sick than when well. For the short-term illnesses considered here, this effect should be negligible because of the opportunity to shift consumption from the period of illness to healthier periods.

Despite the theoretical prediction, nearly all stated-preference studies of WTP for reductions in health risk that test for it find a less-than-proportional relationship between WTP and risk reduction (Hammitt and Graham, 1999). This result suggests inadequate sensitivity to scope, perhaps due to difficulties in communicating the magnitude of risk reduction to survey respondents or to respondents valuing the risk reduction formed by combining their prior estimates of the risk reduction with the value stated in the survey through a Bayesian updating process (Viscusi, 1985, 1989; Hammitt and Graham, 1999). Inadequate sensitivity to risk reduction would be reflected in an estimated value of $\delta < 1$. Conversely, an estimated value of $\delta > 1$ would imply excessive sensitivity to risk reduction, which might occur if the risk is reduced to a level that respondents treat as equivalent to zero (as suggested by prospect theory; Kahneman and Tversky, 1979).

3. Survey

We design and conduct a stated-preference survey to elicit values for reductions in health risks that vary in the baseline probability of illness, reduction in probability, severity and duration of symptoms, and conditional probability of mortality. This section describes the survey instrument and sample.

3.1. Survey Instrument

The survey includes a dichotomous-choice experiment in which respondents decide whether to purchase a safer but more expensive food. The survey instrument is organized as follows. First, respondents are asked about their experience with foodborne illness and their perception of how common it is compared with other health and safety risks. Second, respondents assess their current health using a visual analogue scale (VAS) and the Health Utilities Index Mark 3 (HUI). The VAS is a numbered line with endpoints of 0 and 100 labeled "equivalent to dead" and "perfect health," respectively. The HUI is a generic, preference-based, multiattribute health-status classification system and index that is widely used as a measure of HRQL in clinical studies, population health surveys, and economic evaluation (Feeny et al., 2002). The HUI classifies health according to the degree of function on eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. For each dimension, there are five or six levels of functional impairment that range from complete function to severe impairment. Table 1 shows the HUI classification system.

Third, respondents complete a tutorial designed to help them practice making tradeoffs between the price and safety of food. The tutorial also familiarizes respondents with a visual aid that communicates the probability of risks (Corso et al., 2001). The visual aid contains red and white areas that represent 10,000 meals, where the fraction of the area that is colored red equals the probability of illness.

Fourth, respondents are asked to consider buying food for a meal that only they will eat. Respondents are asked whether they eat a type of food randomly selected from the set {chicken, ground beef, packaged deli meat}. If they do not eat the selected food, respondents are asked about another randomly-selected food. After answering questions about how often they eat the food and how much they typically eat, respondents are presented with a description of the symptoms of foodborne illness associated with the food and are asked to imagine that they experience those symptoms. Respondents assess their health while ill using the VAS and the HUI. Respondents are then told their baseline probability of illness (either 2 in 10,000 or 4 in 10,000 per meal) and informed that they could reduce their risk to 1 in 10,000 per meal by purchasing a safer but more expensive brand of food. The baseline probability of illness and reduction in probability are communicated using the visual aid described above. The risk reduction is described as produced by a stringent safety program established and monitored by the United States Government, that does not use chemicals or irradiation (some respondents might believe that chemicals or radiation would produce health risk). WTP to reduce the probability of illness is elicited using double-bounded, dichotomous-choice questions. Each respondent is asked if he would purchase the safer food if the extra cost per meal were a randomly selected amount from the set {$0.04, $0.10, $0.20, $0.50, $1, $2, and $4}. There is one follow-up question, in which the bid is equal to twice the initial bid if the respondent is willing to pay the initial amount, and equal to half the initial bid otherwise. A sample question is included in the Appendix. Finally, respondents answer follow-up questions about their food-handling practices, acceptance of the hypothetical scenario, and relevant personal characteristics.

Each respondent is asked to value two health-risk reductions that vary in baseline probability of illness, reduction in probability, severity and duration of symptoms, conditional probability of mortality, and type of food affected. Table 2 shows the risk attributes, which are based on the incidence, range of symptoms, and duration of foodborne illness in the United States. Using a full factorial design, the risk attributes are randomly assigned so that each of the possible combinations is asked of some respondents. Respondents who live in a household with at least one child between the ages of 2 and 18 are asked about reducing one risk to their own health and one risk to the health of a randomly-selected child from their household (in random order). Other respondents are asked about reducing two risks to their own health, but are not presented with the same food twice.

3.2. Sample

The survey was fielded to 4,481 randomly-selected adults in the United States. Respondents are members of a demographically representative panel maintained by Knowledge Networks. Households are recruited to the panel using random digital dialing and provided free Internet access and hardware, such as MSN® TV, as a participation incentive. In total, 2,795 interviews were completed in several waves between August and October 2004, yielding a response rate of 62%. We report only results based on risks to the respondent's own health. For respondents asked about reducing risk to a child, we include the response to the question about a risk to the respondent if it was the first valuation question, but exclude it for the 415 respondents who are asked about reducing risks to a child's health before risks to their own health (to protect against order effects). We also exclude 58 respondents who do not eat any of the foods, 18 respondents who declined to answer the WTP questions, and 255 respondents who did not complete the HUI questions or who answered them inconsistently (e.g., rating health in the event of sickness as better than current health). A total of 2,049 respondents and 3,481 risk reductions are included in the analysis.

4. Results

This section reports estimates of how estimated WTP depends on attributes of the health risk, respondent characteristics, and other factors.

4.1. Respondent Characteristics

Table 3 lists the variables used for analysis with the means, standard deviations, and ranges for the sample of 2,049 respondents. Mean age is 46 years, with a range of 18 to 96 years. Forty-seven percent of respondents are male. Seventy-four percent identify themselves as non-Hispanic white, 12 percent as Hispanic, 10 percent as non-Hispanic black, and 4.1 percent as none of these categories. Fifty-two percent of respondents are married, and the mean household size is 2.4 persons. Mean annual household income is about $48,100. Twenty-seven percent of respondents have a college degree. Thirty-nine percent report having contracted foodborne illness. On average, respondents estimate that 32 percent of the United States population contracts foodborne illness each year, a figure that is consistent with an official estimate of 76 million cases per year (Mead et al., 1999). Respondents are significantly more likely to eat chicken and ground beef than packaged deli meat, which results in 40 percent of respondents answering their first question about chicken, 35 percent answering their first question about ground beef, and 25 percent answering their first question about packaged deli meat. On average, respondents report eating the food 5.6 times per month. In follow-up questions related to food preparation, 55 percent of respondents indicate that they are responsible for preparing the food in their households most or all of the time, 62 percent claim to consistently wash their hands while preparing the food, and 67 percent report taking one or more recommended steps to ensure that the food is fully cooked or otherwise safe to eat. In follow-up questions relating to acceptance of the hypothetical scenario, 50, 36, and 14 percent of respondents perceive their risk of foodborne illness to be similar to, smaller than, and larger than that presented in the survey, respectively. Eighty-three percent of respondents express at least some confidence that the safety system does not use chemicals or irradiation (as stated in the survey), and 4 and 6 percent believe the government and private sector are not at all effective in ensuring food safety, respectively.

Table 4 reports the means and standard deviations of the VAS and HUI scores that respondents assign to current health and to the illness, along with the symptom descriptions presented in the survey. The results show general agreement between the two measures of health. Mean VAS score for current health is 0.754, with a range of 0.050 to 1. Mean HUI score for current health is 0.800, with a range of −0.195 to 1. Both VAS scores and HUI scores decrease with the severity of illness. Mean VAS scores for symptom descriptions listed in the table as “mild,” “moderate,” and “severe” are 0.580, 0.461, and 0.417, respectively. The corresponding mean HUI scores are 0.507, 0.260, and 0.105, respectively. Consistent with other comparisons of rating scale and standard gamble estimates (Torrance, 1996), the VAS scores tend to cluster in the middle of the scale while the standard-gamble-based HUI scores are spread out across a wider range. Comparisons between the VAS and HUI scores should be interpreted with caution because possible scores for the VAS range between 0 and 1 while possible scores for the HUI range between -0.359 and 1.

4.2. Effects of Health Risk Attributes on WTP

WTP is modeled as a function of the severity and duration of illness, reduction in its probability, other risk attributes, and demographic and preference characteristics of the respondents, as specified in Equation (2). Regression models are estimated using the maximum-likelihood method under the assumption that WTP follows a lognormal distribution (Alberini, 1995). Results are presented in Table 5.

To examine how WTP depends on the health risk, we estimate a regression that includes only variables for the decrement in HRQL associated with illness (as measured by the HUI), the duration of illness, and the reduction in its probability.

Model 1 indicates that WTP increases with the severity and duration of the illness to be prevented. We reject the hypotheses that WTP is insensitive to the HRQL decrement ($p = 0.002$) and duration of illness ($p = 0.039$). However, the estimated coefficient for the loss in HRQL is significantly smaller than one, which allows us to reject the hypothesis that WTP is proportional to changes in HRQL ($p < 0.001$). Similarly, the estimated coefficient for the duration of illness is significantly smaller than one, which allows us to reject the hypothesis that WTP is proportional to the duration of illness ($p < 0.001$). Taken together, the estimated coefficients for the loss in HRQL and duration of illness suggest diminishing marginal WTP for changes in HRQL and duration of illness and that WTP to reduce the risk of illness varies less than proportionally to changes in expected QALYs. As described in Section 2.1, previous studies also found that marginal WTP decreases with the duration of the illness to be prevented, but increased with illness severity.

The estimated coefficient for the risk reduction allows us to reject the hypothesis that WTP is insensitive to the probability of risk reduction ($p < 0.001$). Moreover, the estimated coefficient is significantly smaller than one, suggesting that WTP is less than proportional to the reduction in probability of illness ($p < 0.001$). These results are consistent with most previous studies that use stated-preference methods to estimate the value of reducing health risk (Hammitt and Graham, 1999). However, single-bounded models estimated using only responses to the initial bid suggest that estimated WTP is nearly proportional to the probability of risk reduction (see Section 4.5 below).

The estimated less-than-proportional relationship between WTP and change in expected QALYs might be explained by inadequate sensitivity of estimated WTP to the severity and duration of illness, as an alternative to diminishing marginal WTP for severity and duration. However, this hypothesis seems inconsistent with our finding that WTP is more sensitive to the change in probability of illness than to illness severity and duration. In every model, the estimated coefficient for risk reduction is substantially larger than the estimated coefficients for the HRQL decrement and duration of illness. A large body of psychological research shows that individuals have difficulty comprehending small probabilities (Kahneman and Tversky, 1973; Baron, 1997), which could explain inadequate sensitivity to probability. It seems implausible that respondents would be less sensitive to differences in duration (of 1, 3, and 7 days) and differences in severity (which they spent time evaluating using both VAS and HUI instruments) than to small changes in probability.

4.3. Effects of Respondent Characteristics on WTP

Model 2 adds respondents' personal characteristics to the variables included in Model 1. The additional variables have little effect on the estimated coefficients of the health risk attributes, which is expected given the randomized study design. The signs, magnitudes, and standard errors of the estimated coefficients in Model 2 remain largely unchanged from Model 1.

Estimated WTP to reduce the risk of illness is significantly associated with some of the personal characteristics. Estimated WTP increases with age at a rate of 1.2 percent per year. Males are estimated to value risk reduction 41 percent less than females. Non-Hispanic blacks and Hispanics are estimated to value risk reduction about 3 and 2 times more, respectively, than non-Hispanic whites. Married respondents are estimated to value risk reduction 22 percent less than respondents who are not married, with estimated WTP increasing at a rate of 11 percent per additional household member. Respondents with a college degree are estimated to be willing to pay 26 percent less than respondents with less education. The estimated coefficient for household income is small, negative, and not statistically significant ($p = 0.428$), potentially reflecting the low cost of the risk reduction.

4.4. Effects of Risk Perception, Acceptance of Hypothetical Scenario, and Other Factors on WTP

Model 3 adds responses to follow-up questions, other risk attributes, and interactions between question order and the type of food to the model. The additional variables have little effect on the estimated coefficients of the health risk attributes and of most respondent characteristics. However, the magnitudes of the coefficient estimates for males and non-Hispanic blacks are smaller than in model 2. Males are estimated to value the risk reduction 30 percent less than females, and non-Hispanic blacks are estimated to value the risk reduction 2.2 times more than non-Hispanic whites.

Estimated WTP to reduce the risk of illness is significantly associated with risk perception and prior experience with foodborne illness. Respondents who perceive higher prevalence of foodborne illness in the United States are estimated to value the risk reduction more than respondents who perceive lower prevalence, with estimated WTP increasing at a rate of 1.4 percent per percentage-point increase in perceived prevalence. Respondents who perceive their risk of foodborne illness to be higher or lower than that presented to them in the survey are estimated to value the risk reduction 39 percent more and 25 percent less, respectively, than respondents who perceive their risk to be similar to the stated magnitude. Respondents who report having contracted foodborne illness are estimated to value the risk reduction 18 percent less than other respondents.

Estimated WTP to reduce the risk of illness is also significantly associated with acceptance of the hypothetical scenario. Respondents who are not confident that the safety system does not use chemicals or radiation (as stated in the survey) are estimated to value the risk reduction 63 percent less than respondents who express at least some confidence. Respondents who believe that the government and private sector are not at all effective in ensuring food safety are estimated to value the risk reduction 42 percent less and 88 percent more, respectively, than respondents who hold more favorable views. These results suggest that respondents are willing to pay more for safety when they believe that government regulation of the private sector is effective and necessary.

The type of food presenting the risk, the frequency with which the respondent eats the food, and protective behaviors related to food preparation are significantly associated with estimated WTP. Respondents are estimated to value reducing risk associated with chicken more than they value risks associated with ground beef and packaged deli meat, but the estimates depend on question order. When comparing responses to the first risk reduction, respondents are estimated to value reducing a risk of illness associated with chicken 76 percent more than the value associated with ground beef and 54 percent more than the value associated with packaged deli meat. We find no statistically significant differences between estimated WTP to reduce the risk of illness associated with ground beef and packaged deli meat. Question order has an effect on the estimated coefficients for ground beef and packaged deli meat, but not on the estimated coefficient for chicken. Respondents are estimated to value reducing risk associated with ground beef 64 percent more it is the second risk reduction as opposed to the first, and packaged deli meat 43 percent more if it is the second risk reduction as opposed to the first. Estimated WTP decreases with the monthly frequency of consumption at a rate of 1.7 percent per additional meal. Respondents who are more likely to be the food preparer in their household are estimated to value the risk reduction less than others, with estimated WTP decreasing at a rate of 8.1 percent per point on a five-point scale. Respondents who claim to consistently wash their hands while preparing the food are estimated to value the risk reduction 41 percent more than others. Respondents who report taking one or more recommended steps to ensure that the food is fully cooked or otherwise safe to eat are estimated to value the risk reduction 33 percent more than respondents who do not engage in such self-protective behaviors.

Model 3 includes dummy variables for the conditional mortality risks. The estimated coefficients for the conditional mortality risks of 1 in 1,000 and 1 in 10,000 are small, positive, and statistically insignificant ($p = 0.934$ and $p = 0.768$, respectively). The lack of sensitivity to conditional mortality risk may be explained by the small reduction in absolute mortality risk, which ranges from only 1 in 100 million to 30 in 100 million. The point estimates of WTP to reduce mortality risk, which range from 1 to 6 cents per meal, imply an average value per statistical life of $1.3 million, which is comparable to recommended values of $4 million to 9 million (Viscusi and Aldy, 2003), $1.5 million to 2.5 million (Mrozek and Taylor, 2002), and $1.5 million to 4.8 million (Alberini et al., 2004).

To explore for potential effects of question order, we estimated Model 4 using only responses to the first risk reduction and Model 5 using only responses to the second risk reduction. The estimated coefficients have the same signs and similar magnitudes in both models, suggesting that question order has little effect on our results.

4.5. Value per Statistical Case of Foodborne Illness

Table 6 reports estimates of the value per statistical case of foodborne illness, stratified by the severity and duration of illness. The estimates are calculated using the corresponding regression models in Table 5 to predict median WTP for the mean respondent for each risk reduction, dividing by the risk reduction, and then taking the average of these values. For Model 3, the estimated values range from $9,200 for mild symptoms that last one day to $13,600 for severe symptoms that last seven days. The small proportional difference between these values reflects the less-than-proportional relationships between estimated WTP and both duration and severity of illness.

The average value per QALY is reported in the last column of Table 6, where the range corresponds to the range of WTP values predicted by the five models. The average WTP per QALY ranges from $501,000 to $1.04 million, which is somewhat larger than the median value of $265,345 (in 1997 dollars) reported in a review of the value per statistical life literature that calculated implied estimates of WTP per QALY by assuming that WTP is proportional to QALYs gained (Hirth et al., 2000). Because WTP varies less than proportionately with the decrement in HRQL and duration of illness, the average WTP per QALY is smaller for the cases with more severe symptoms and longer durations.

4.6. Sensitivity Analysis

Although regression models of WTP based on double-bounded, dichotomous-choice questions may produce more efficient estimates than those obtained using the single-bounded format (Hanemann et al., 1991; Alberini, 1995), the initial bid may influence responses to the follow-up questions yielding biased estimates (Alberini et al., 1997). To investigate the magnitude of any follow-up effect, we estimate the regression models in Table 5 using only the (single-bounded) responses to the initial bid. The results are shown in Table 7. The single- and double-bounded estimates are similar, with one important exception. The estimated coefficient for the risk reduction is substantially larger in the single-bounded models, suggesting that estimated WTP may be proportional to the probability of risk reduction. Using single-bounded estimates, we reject the hypothesis that WTP is insensitive to risk reduction, but cannot reject the hypothesis that WTP is proportional to the risk reduction in three of the five regression models. For example, the estimated coefficient of the risk reduction in Model 3 is 0.821 as opposed to 0.547 in the double-bounded model and the p-value for rejecting the hypothesis of proportionality is 0.168. For Model 4, the coefficient estimated using the single-bounded method is 0.883 as opposed to 0.500 in the double-bounded model and the p-value for rejecting the proportionality hypothesis is 0.523. Using the single-bounded estimates, we can reject the hypothesis that WTP is proportional to the probability of risk reduction in Model 1 ($p = 0.049$) and Model 2 ($p = 0.057$), but not in Model 5 ($p = 0.132$).

5. Conclusions

In a survey of the general United States population, we find that WTP to reduce the risk of illness increases with, but is less than proportional to, changes in HRQL and duration of illness.

We find that respondents exhibit diminishing marginal WTP for severity and duration of illness avoided. Previous studies also found that marginal WTP decreases with duration of the illness avoided (e.g., Tolley et al., 1994; Johnson et al., 1997; Van Houtven et al., 2003), but some studies found that marginal WTP increases with severity (Johnson et al., 1997; Van Houtven et al., 2003). Our results suggest that individuals do not have a constant WTP for changes in QALYs. Further research should examine whether these results may be generalized to illnesses of longer duration.

This study provides empirical support for the claim that cost-effectiveness analyses using QALYs are not consistent with benefit-cost analysis (Hammitt, 2002a). In particular, economic analyses that rely on a single WTP per QALY or corresponding cost per QALY threshold are inconsistent with economic welfare theory. CEA is not equivalent to BCA and its use must be justified on extra-welfarist or other principles (Culyer, 1989, 1990).

Table 1. Health Utilities Index Mark III

Attribute	Level
Vision	1. Able to see well enough to read ordinary newsprint and recognize a friend on the other side of the street, without glasses or contact lenses. 2. Able to see well enough to read ordinary newsprint and recognize a friend on the other side of the street, but with glasses. 3. Able to read ordinary newsprint with or without glasses but unable to recognize a friend on the other side of the street, even with glasses. 4. Able to recognize a friend on the other side of the street with or without glasses but unable to read ordinary newsprint, even with glasses. 5. Unable to read ordinary newsprint and unable to recognize a friend on the other side of the street, even with glasses. 6. Unable to see at all.
Hearing	1. Able to hear what is said in a group conversation with at least three other people, without a hearing aid. 2. Able to hear what is said in a conversation with one other person in a quiet room without a hearing aid, but requires a hearing aid to hear what is said in a group conversation with at least three other people. 3. Able to hear what is said in a conversation with one other person in a quiet room with a hearing aid, and able to hear what is said in a group conversation with at least three other people, with a hearing aid. 4. Able to hear what is said in a conversation with one other person in a quiet room, without a hearing aid, but unable to hear what is said in a group conversation with at least three other people even with a hearing aid. 5. Able to hear what is said in a conversation with one other person in a quiet room with a hearing aid, but unable to hear what is said in a group conversation with at least three other people even with a hearing aid. 6. Unable to hear at all.
Speech	1. Able to be understood completely when speaking with strangers or friends. 2. Able to be understood partially when speaking with strangers but able to be understood completely when speaking with people who know me well. 3. Able to be understood partially when speaking with strangers or people who know me well. 4. Unable to be understood when speaking with strangers but able to be understood partially by people who know me well. 5. Unable to be understood when speaking to other people (or unable to speak at all).
Ambulation	1. Able to walk around the neighborhood without difficulty, and without walking equipment. 2. Able to walk around the neighborhood with difficulty; but does not require walking equipment or the help of another person. 3. Able to walk around the neighborhood with walking equipment, but without the help of another person. 4. Able to walk only short distances with walking equipment, and requires a wheelchair to get around the neighborhood. 5. Unable to walk alone, even with walking equipment. Able to walk short distances with the help of another person, and requires a wheelchair to get around the neighborhood. 6. Cannot walk at all.
Dexterity	1. Full use of two hands and ten fingers. 2. Limitations in the use of hands or fingers, but does not require special tools or help of another person. 3. Limitations in the use of hands or fingers, is independent with use of special tools (does not require the help of another person). 4. Limitations in the use of hands or fingers, requires the help of another person for some tasks (not independent even with use of special tools). 5. Limitations in use of hands or fingers, requires the help of another person for most tasks (not independent even with use of special tools). 6. Limitations in use of hands or fingers, requires the help of another person for all tasks (not independent even with use of special tools).
Emotion	1. Happy and interested in life. 2. Somewhat happy. 3. Somewhat unhappy. 4. Very unhappy. 5. So unhappy that life is not worthwhile.
Cognition	1. Able to remember most things, think clearly and solve day to day problems. 2. Able to remember most things, but have a little difficulty when trying to think and solve day to day problems. 3. Somewhat forgetful, but able to think clearly and solve day to day problems. 4. Somewhat forgetful, and have a little difficulty when trying to think or solve day to day problems. 5. Very forgetful, and have great difficulty when trying to think or solve day to day problems. 6. Unable to remember anything at all
Pain	1. Free of pain and discomfort. 2. Mild to moderate pain that prevents no activities. 3. Moderate pain that prevents a few activities. 4. Moderate to severe pain that prevents some activities. 5. Severe pain that prevents most activities.

Table 2. Risk Attributes (Full-Factorial Design)

Risk Reduction	Severity of Symptoms	Duration of Illness	Conditional Mortality	Type of Food
1 in 10,000	Mild	1 day	None	Chicken
3 in 10,000	Moderate	3 days	1 in 10,000	Ground Beef
	Severe	7 days	1 in 1,000	Packaged Deli Meat

Table 3. Variables and Descriptive Statistics

Variable	Mean	Std. Dev.	Range
Log(Risk Reduction)	-8.641	0.549	[-9.210, -8.112]
Log(Loss in HUI Score)	-0.897	0.790	[-4.500, 0.252]
Log(Duration of Illness)	0.971	0.794	[0 , 1.946]
Conditional Mortality Risk			
1 in 1,000	0.317	0.465	[0, 1]
1 in 10,000	0.348	0.477	[0, 1]
None	0.335	0.472	[0, 1]
Age	45.994	16.717	[18, 96]
Male	0.473	0.499	[0, 1]
Race and Ethnicity			
White, Non-Hispanic	0.740	0.439	[0, 1]
Hispanic	0.119	0.323	[0, 1]
Black, Non-Hispanic	0.101	0.301	[0, 1]
Other, Non-Hispanic	0.041	0.198	[0, 1]
Married	0.523	0.500	[0, 1]
Household Size	2.447	1.306	[1, 9]
Log(Household Income)	10.463	0.914	[7.247, 12.206]
College Degree	0.266	0.442	[0, 1]
Prior Foodborne Illness	0.388	0.487	[0, 1]
Perceived Prevalence of Foodborne Illness	31.878	22.954	[0, 100]
Perception of Own Risk Versus Stated Risk			
Perceive Own Risk to be Higher	0.140	0.347	[0, 1]
Perceive Own Risk to be the Same	0.502	0.500	[0, 1]
Perceive Own Risk to be Lower	0.358	0.480	[0, 1]
Confidence in Safety System			
Not Confident	0.354	0.478	[0, 1]
Somewhat Confident	0.481	0.500	[0, 1]
Very Confident	0.165	0.371	[0, 1]
Low Trust in Government	0.040	0.196	[0, 1]
Low Trust in Private Sector	0.060	0.238	[0, 1]
Type of Food			
Chicken	0.403	0.491	[0, 1]
Ground Beef	0.345	0.475	[0, 1]
Packaged Deli Meat	0.252	0.434	[0, 1]
Monthly Frequency of Consumption	5.606	6.284	[1, 30]
Degree of Responsibility for Preparing Food	2.394	1.481	[0, 4]
Wash Hands	0.616	0.487	[0, 1]
Safe Food Practices	0.673	0.469	[0, 1]
Second Risk	0.411	0.492	[0, 1]
Second Risk is Chicken	0.152	0.359	[0, 1]
Second Risk is Ground Beef	0.152	0.359	[0, 1]
Second Risk is Packaged Deli Meat	0.108	0.310	[0, 1]

Note: Statistics are based on the first risk presented to the 2,049 respondents, with the exception of variables relating to question order. A total of 1,432 respondents were also asked about a second risk.

Table 4. VAS and HUI Scores Assigned to Current Health and to Symptoms of Illness

Health State	VAS Score		HUI Score	
	Mean	Std. Dev.	Mean	Std. Dev.
Current Health	0.754	0.167	0.800	0.204
Mild Symptoms You will have an upset stomach and will feel tired, but these symptoms will not prevent you from going to work or from doing most of your regular activities.	0.580	0.207	0.507	0.268
Moderate Symptoms You will have an upset stomach, fever, and will need to lie down most of the time. You will be tired and will not feel like eating or drinking much. Occasionally, you will have painful cramps in your stomach. In addition, you will have some diarrhea and will need to stay close to a bathroom. While you are sick, you will not be able to go to work or do most of your regular activities.	0.461	0.228	0.260	0.304
Severe Symptoms You will have to be admitted to a hospital. You will have painful cramps in your stomach, fever, and will need to spend most of your time lying in bed. You will need to vomit and will have severe diarrhea that will leave you seriously dehydrated. Because you will be unable to eat or drink much, you will need to have intravenous tubes put in your arm to provide nourishment.	0.417	0.249	0.105	0.300

Note: Possible scores for the VAS range between 0 and 1, while possible scores for the HUI range between -0.359 and 1

Table 5. Regression Results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	4.798***	4.854***	4.722***	4.412***	5.037***
	(0.687)	(0.879)	(0.887)	(1.140)	(1.407)
Log(Risk Reduction)	0.482***	0.491***	0.547***	0.500***	0.607***
	(0.078)	(0.077)	(0.075)	(0.097)	(0.120)
Log(Loss in HUI Score)	0.170***	0.200***	0.191***	0.156**	0.219***
	(0.054)	(0.054)	(0.052)	(0.069)	(0.081)
Log(Duration of Illness)	0.110**	0.125**	0.113**	0.092	0.142*
	(0.053)	(0.053)	(0.051)	(0.067)	(0.081)
Conditional Mortality Risk of 1 in 1,000			0.008	-0.046	0.094
			(0.101)	(0.131)	(0.160)
Conditional Mortality Risk of 1 in 10,000			0.029	-0.062	0.151
			(0.100)	(0.128)	(0.159)
Age		0.011***	0.013***	0.014***	0.011**
		(0.003)	(0.003)	(0.004)	(0.004)
Male		-0.532***	-0.358***	-0.392***	-0.313**
		(0.085)	(0.089)	(0.116)	(0.140)
Black, Non-Hispanic		1.077***	0.800***	0.806***	0.783***
		(0.156)	(0.154)	(0.194)	(0.252)
Hispanic		0.712***	0.668***	0.775***	0.504**
		(0.142)	(0.138)	(0.175)	(0.225)
Other Race, Non-Hispanic		-0.308	-0.293	-0.382	-0.105
		(0.218)	(0.212)	(0.261)	(0.360)
Married		-0.254***	-0.312***	-0.200*	-0.466***
		(0.096)	(0.094)	(0.121)	(0.151)
Household Size		0.102***	0.081**	0.066	0.101
		(0.039)	(0.039)	(0.047)	(0.068)
Log(Household Income)		-0.041	0.006	-0.017	0.043
		(0.051)	(0.050)	(0.065)	(0.079)
College Degree		-0.306***	-0.234**	-0.265**	-0.190
		(0.099)	(0.096)	(0.126)	(0.149)
Prior Foodborne Illness			-0.204**	-0.198*	-0.239*
			(0.087)	(0.112)	(0.139)
Perceived Prevalence			0.014***	0.012***	0.017***
			(0.002)	(0.003)	(0.003)
Perceive Own Risk to be Higher			0.332***	0.301*	0.352*
			(0.127)	(0.168)	(0.193)
Perceive Own Risk to be Lower			-0.287***	-0.286**	-0.310**
			(0.091)	(0.118)	(0.144)
Not Confident in Safety System			-0.994***	-0.876***	-1.158***
			(0.123)	(0.157)	(0.195)
Somewhat Confident in Safety System			0.005	0.046	-0.055
			(0.093)	(0.119)	(0.147)

Continued on next page

Table 5. Regression Results (continued)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Low Trust in Government			-0.546** (0.223)	-0.671** (0.286)	-0.325 (0.356)
Low Trust in Private Sector			0.632*** (0.199)	0.582** (0.257)	0.679** (0.312)
Ground Beef			-0.562*** (0.127)	-0.556*** (0.128)	-0.011 (0.154)
Packaged Deli Meat			-0.432*** (0.161)	-0.354** (0.176)	-0.203 (0.227)
Monthly Frequency of Consumption			-0.017** (0.007)	-0.018** (0.009)	-0.015 (0.011)
Responsibility for Preparing Meals			-0.084** (0.033)	-0.085** (0.042)	-0.094* (0.053)
Wash Hands			0.340*** (0.085)	0.401*** (0.110)	0.283** (0.134)
Safe Food Practices			0.288** (0.126)	0.410** (0.161)	0.089 (0.205)
Second Risk			-0.041 (0.136)		
Second Risk is Ground Beef			0.535*** (0.195)		
Second Risk is Packaged Deli Meat			0.397* (0.211)		
Sigma	2.082 (0.052)	2.029 (0.051)	1.941 (0.048)	1.930 (0.062)	1.937 (0.077)
Sample Size	3,481	3,481	3,481	2,049	1,432
Log Likelihood	-3,865.4	-3,784.1	-3,658.7	-2,191.1	-1,460.5

Note: Standard errors are in parentheses. ***, **, and * denote statistical significance from $\beta = 0$ at 1, 5, and 10 percent, respectively, based on likelihood-ratio tests.

Table 6. Estimated Value per Statistical Case of Foodborne Illness and Average WTP per QALY (US$)

Severity of Symptoms		Duration of Illness	Model 1	Model 2	Model 3	Model 4	Model 5	WTP per QALY	
Description	Loss in HUI Score							Low	High
Mild	0.293	1 day	$9,100	$9,100	$9,200	$8,500	$10,300	$790,000	$953,000
Mild	0.293	3 days	$10,200	$10,400	$10,500	$9,400	$12,000	$761,000	$970,000
Mild	0.293	7 days	$11,200	$11,600	$11,500	$10,200	$13,600	$654,000	$869,000
Moderate	0.540	1 day	$10,100	$10,300	$10,400	$9,400	$11,800	$818,000	$1,025,000
Moderate	0.540	3 days	$11,400	$11,800	$11,800	$10,400	$13,700	$720,000	$953,000
Moderate	0.540	7 days	$12,500	$13,100	$12,900	$11,200	$15,500	$552,000	$762,000
Severe	0.695	1 day	$10,500	$10,800	$10,900	$9,800	$12,400	$821,000	$1,045,000
Severe	0.695	3 days	$11,900	$12,400	$12,300	$10,800	$14,500	$688,000	$925,000
Severe	0.695	7 days	$13,000	$13,800	$13,600	$11,700	$16,400	$501,000	$703,000

Note: Estimates of WTP per QALY are based on the range of WTP across models.

Table 7. Regression Results Based on Single-Bounded Estimates of WTP

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	7.573*** (1.183)	7.470*** (1.489)	7.017*** (1.502)	7.163*** (2.087)	6.380*** (2.139)
Log(Risk Reduction)	0.741*** (0.132)	0.747*** (0.131)	0.821*** (0.130)	0.883*** (0.183)	0.724*** (0.184)
Log(Loss in HUI Score)	0.207** (0.087)	0.255*** (0.088)	0.248*** (0.086)	0.238* (0.123)	0.234** (0.119)
Log(Duration of Illness)	0.139 (0.086)	0.158* (0.086)	0.132 (0.084)	0.099 (0.118)	0.159 (0.118)
Conditional Mortality Risk of 1 in 1,000			-0.144 (0.166)	-0.138 (0.232)	-0.152 (0.236)
Conditional Mortality Risk of 1 in 10,000			0.005 (0.164)	0.017 (0.228)	-0.012 (0.235)
Age		0.021*** (0.005)	0.024*** (0.005)	0.030*** (0.007)	0.018*** (0.007)
Male		-0.768*** (0.142)	-0.477*** (0.148)	-0.580*** (0.209)	-0.385* (0.207)
Black, Non-Hispanic		1.433*** (0.264)	1.001*** (0.257)	1.132*** (0.354)	0.810** (0.372)
Hispanic		0.883*** (0.232)	0.824*** (0.227)	1.035*** (0.315)	0.563* (0.331)
Other Race, Non-Hispanic		-0.139 (0.354)	-0.119 (0.348)	-0.096 (0.467)	-0.136 (0.526)
Married		-0.323** (0.156)	-0.397*** (0.154)	-0.222 (0.213)	-0.621*** (0.225)
Household Size		0.142** (0.064)	0.112* (0.064)	0.077 (0.084)	0.175* (0.103)
Log(Household Income)		-0.066 (0.083)	0.014 (0.082)	-0.008 (0.115)	0.050 (0.117)
College Degree		-0.495*** (0.162)	-0.362** (0.158)	-0.415* (0.223)	-0.270 (0.220)
Prior Foodborne Illness			-0.177 (0.143)	-0.169 (0.200)	-0.203 (0.206)
Perceived Prevalence			0.023*** (0.004)	0.024*** (0.005)	0.022*** (0.005)
Perceive Own Risk to be Higher			0.489** (0.213)	0.707** (0.313)	0.250 (0.287)
Perceive Own Risk to be Lower			-0.467*** (0.150)	-0.492** (0.209)	-0.430** (0.213)
Not Confident in Safety System			-1.483*** (0.214)	-1.428*** (0.296)	-1.550*** (0.308)
Somewhat Confident in Safety System			0.114 (0.151)	0.291 (0.211)	-0.086 (0.216)

Continued on next page

Table 7. Regression Results Based on Single-Bounded Estimates of WTP (continued)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Low Trust in Government			-0.383 (0.364)	-0.590 (0.506)	-0.060 (0.523)
Low Trust in Private Sector			0.658** (0.322)	0.658 (0.457)	0.615 (0.450)
Ground Beef			-0.826*** (0.212)	-0.870*** (0.236)	0.080 (0.226)
Packaged Deli Meat			-0.723*** (0.269)	-0.510 (0.317)	-0.426 (0.334)
Monthly Frequency of Consumption			-0.025** (0.011)	-0.023 (0.016)	-0.024 (0.016)
Responsibility for Preparing Meals			-0.114** (0.054)	-0.142* (0.075)	-0.104 (0.078)
Wash Hands			0.499*** (0.140)	0.702*** (0.200)	0.302 (0.197)
Safe Food Practices			0.319 (0.208)	0.727** (0.292)	-0.218 (0.305)
Second Risk			-0.214 (0.224)		
Second Risk is Ground Beef			0.914*** (0.323)		
Second Risk is Packaged Deli Meat			0.684** (0.347)		
Sigma	2.988 (0.184)	2.929 (0.180)	2.794 (0.170)	3.008 (0.255)	2.507 (0.220)
Sample Size	3,481	3,481	3,481	2,049	1,432
Log Likelihood	-2,014.8	-1,951.5	-1,843.0	-1,095.0	-738.4

Note: Standard errors are in parentheses. ***, **, and * denote statistical significance from $\beta = 0$ at 1, 5, and 10 percent, respectively, based on likelihood ratio tests.

Appendix

The table below summarizes the differences between the Superior Safety System chicken and the standard chicken. Please consider which type of chicken you would buy for a meal that only you would eat. Remember that the extra money you spend for a meal with Superior Safety System chicken is money that you could no longer spend on other things you might want or need.

Type of Chicken	Your Chance of Illness per Meal	Cost per Meal
Standard	4 in 10,000	standard cost
Superior Safety System	1 in 10,000	$0.50 more per meal than standard chicken

Whether you eat the Superior Safety System chicken or the standard chicken, if you get sick:

- You will have an upset stomach, fever, and will need to lie down most of the time. You will be tired and will not feel like eating or drinking much. Occasionally, you will have painful cramps in your stomach. In addition, you will have some diarrhea and will need to stay close to a bathroom. While you are sick, you will not be able to go to work or do most of your regular activities.
- You will have these symptoms for 3 days.
- There is 1 in 10,000 chance that you will die from this sickness.

If Superior Safety System chicken cost $0.50 more per meal than standard chicken, which type of chicken would you purchase?

○ Standard

○ Superior Safety System

Acknowledgement

This work was supported by the Economic Research Service of the United States Department of Agriculture. We thank Joni Hersch, Nicolas Treich, Kip Viscusi, and Milt Weinstein for helpful comments.

References

Alberini, A., "Efficiency vs. Bias of Willingness-to-Pay Estimates: Bivariate and Interval-Data Models," *Journal of Environmental Economics and Management* 29: 169-180, 1995.

Alberini, A., M. Cropper, A. Krupnick, and N. Simon, "Does the Value of a Statistical Life Vary with Age and Health Status? Evidence from the US and Canada," *Journal of Environmental Economics and Management* 48: 769-792, 2004.

Alberini, A., "Efficiency vs. Bias of Willingness-to-Pay Estimates: Bivariate and Interval-Data Models," *Journal of Environmental Economics and Management* 29: 169-180, 1995.

Bala, M.V., L.L. Wood, G.A. Zarkin, E.C. Norton, A. Gafni, and B. O'Brien, "Valuing Outcomes in Health Care: A Comparison of Willingness to Pay and Quality-Adjusted Life Years," *Journal of Clinical Epidemiology* 51: 667-676, 1998.

Baron, J., "Biases in the Quantitative Measurement of Values for Public Decisions," *Psychological Bulletin* 122: 72-88, 1997.

Bleichrodt, H., P. Wakker, and M. Johannesson, "Characterizing QALYs by Risk Neutrality," *Journal of Risk and Uncertainty* 15: 107-114, 1997.

Corso, P.S., J.K. Hammitt, and J.D. Graham, "Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation," *Journal of Risk and Uncertainty* 23(2): 165-184, 2001.

Culyer, A.J., "The Normative Economics of Health Care Finance and Provision," *Oxford Review of Economic Policy* 5: 34-58, 1989.

Culyer, A.J., "Commodities, Characteristics of Commodities, Characteristics of People, Utilities, and the Quality of Life," in S. Baldwin, C. Godfrey, and C. Propper, eds., *Quality of Life: Perspectives and Policies*, Routledge, London, 1990.

EuroQol Group, "EuroQoL – A New Facility for the Measurement of Health-Related Quality of Life," *Health Policy* 16(3): 199-208, 1990.

Feeny, D., W. Furlong, G.W. Torrance, C.H. Goldsmith, Z. Zhu, S. DePauw, M. Denton, and M. Boyle, "Multiattribute and Single-Attribute Utility Functions for the Health Utilities Index Mark 3 System," *Medical Care* 40: 113-128, 2002.

Garber, A.M., and C.E. Phelps, "Economic Foundations of Cost-Effectiveness Analysis," *Journal of Health Economics* 16: 1-31, 1997.

Gyrd-Hansen, D., "Willingness to Pay for a QALY," *Health Economics* 12: 1049-1060, 2003.

Hammitt, J.K. "Evaluating Contingent Valuation of Environmental Health Risks: The Proportionality Test," *Association of Environmental and Resource Economists Newsletter* 20(1): 14-19, May 2000. Reprinted in *Stated Preference: What Do We Know? Where Do We Go? (Proceedings)*, Report number EE-0436, U.S. Environmental Protection Agency, October 2000.

Hammitt, J.K., "QALYs versus WTP," *Risk Analysis* 22(5): 985-1001, 2002a.

Hammitt, J.K., "How Much is a QALY Worth? Admissible Utility Functions for Health and Wealth" Harvard Center for Risk Analysis, May 2002b.

Hammitt, J.K., and J.D. Graham, "Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?" *Journal of Risk and Uncertainty* 18(1): 33-62, 1999.

Hanemann, M.W., "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review* 81: 635-647, 1991.

Hanemann, W.M., J. Loomis, and B. Kanninen, "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation," *American Journal of Agricultural Economics* 73: 1255-1261, 1991.

Hirth, R.A., M.E. Chernew, E. Miller, A.M. Fendrick, and W.G. Weissert, "Willingness to Pay for a Quality-Adjusted Life Year: In Search of a Standard," *Medical Decision Making* 20: 332-342, 2000.

Johannesson, M., "The Relationship between Cost-Effectiveness Analysis and Cost-Benefit Analysis," *Social Science and Medicine* 41: 483-489, 1995.

Johnson, F.R., M.R. Banzhaf, and W.H. Desvousges, "Willingness to Pay for Improved Respiratory and Cardiovascular Health: A Multiple-Format, Stated-Preference Approach," *Health Economics* 9: 295-317, 2000.

Johnson F.R., E.E. Fries, and H.S. Banzhaf, "Valuing morbidity: an integration of the willingness-to-pay and health-status index literatures," *Journal of Health Economics* 16: 641-665, 1997.

Jones-Lee, M.W., G. Loomes, and P.R. Philips, "Valuing the Prevention of Non-Fatal Road Injuries: Contingent Valuation vs. Standard Gambles," *Oxford Economic Papers* 47: 676-695, 1995.

Kahneman, D., and A. Tversky, "On the Psychology of Prediction," *Psychology Review* 80: 237-251, 1973.

Kahneman, D., and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47: 263-291, 1979.

Kaplan, R.M., J.P. Anderson, and T.G. Ganiats, "The Quality of Well-Being Scale: Rationale for a Single Quality of Life Index," in Walker, S.R., and R.M. Rosser (eds.), *Quality of Life Assessment: Key Issues in the 1990s*, Kluwer, Dordrecht, 1993.

Krabbe, P.F.M., M. Essink-Bot, and G.J. Bonsel, "The Comparability and Reliability of Five Health-State Valuation Methods. *Social Science and Medicine* 45: 1641-1652, 1997.

Mead, P.S., L. Slutsker, V. Dietz, L.F. McCaig, J.S. Bresee, C. Shapiro, P.M. Griffin, and R.V. Tauxe, "Food-Related Illness and Death in the United States," *Emerging Infectious Diseases* 5: 840-842, 1999.

Mrozek, J.R., and L.O. Taylor, "What Determines the Value of Life? A Meta-Analysis," *Journal of Policy Analysis and Management* 21: 253–270, 2002.

O'Brien, B., and J.L. Viramontes, "Willingness to Pay: A Valid and Reliable Measure of Health State Preference?" *Medical Decision Making* 14: 289-297, 1994.

Pliskin, J.S., D.S. Shepard, and M.C. Weinstein, "Utility Functions for Life Years and Health Status," *Operations Research* 28: 206-224, 1980.

Pratt, J.W., and R.J. Zeckhauser, "Willingness to Pay and the Distribution of Risk and Wealth," Journal of Political Economy 104: 747-763, 1996.

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

Torrance, G.W, D.H. Feeny, W.J. Furlong, R.D. Barr, Y. Zhang, Q. Wang, "Multiattribute Utility Function for a Comprehensive Health Status Classification System: Health Utilities Index Mark 2," *Medical Care* 34: 702–722, 1996.

Tversky, A., and D. Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5: 297-323, 1992.

Van Houtven, G., M. Rousu, J.-C. Yang, C. Pringle, W. Wagstaff, and J. DePlatchett, *Valuation of Morbidity Losses: Meta-Analysis of Willingness-to-Pay and Health Status Measures*, Research Triangle Institute, 2003.

Viscusi, W.K., "A Bayesian Perspective on Biases in Risk Perception," *Economic Letters* 17: 59-62, 1985.

Viscusi, W.K, "Prospective Reference Theory: Toward an Explanation of the Paradoxes," *Journal of Risk and Uncertainty* 2: 235-263, 1989.

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