Quality adjusted life years (QALYs) and willingness to pay (WTP) are alternative measures of the value of reductions in health risk that are often used in evaluating environmental, health, and safety practices. Although both methods are based on individual preferences, the underlying assumptions differ. The different bases yield systematically different conclusions about the relative value of reducing health and mortality risks to individuals that differ in age, preexisting health conditions, income, and other factors. The choice of which method to use depends on judgments about what constraints should be placed on individual preferences and what factors should be considered in aggregating preferences across people.

KEY WORDS: Health metric; valuation; quality adjusted life years; willingness to pay

1. INTRODUCTION

Comparative risk analysis and life cycle impact assessment attempt to characterize the portfolio of health risks associated with a specific policy or product. To determine whether the net effect of a policy on health is positive or negative, it is often necessary to aggregate disparate health risks, which may be fatal or nonfatal, cause cancer or other disease, and vary in other characteristics. For example, comparative risk analysis of drinking-water-treatment systems requires a method for comparing risks of gastrointestinal illness caused by microbial pathogens with risks of cancer caused by disinfection byproducts. Life cycle assessment of the effects of coal-burning electricity plants requires aggregating fatal and nonfatal injuries in coal mining and transport with respiratory illness and fatalities from particulate air pollution.

A number of alternative approaches to valuing changes in health and mortality risks have been developed. Two of the most prominent are the “quality adjusted life year” (QALY) and “willingness to pay” (WTP) frameworks. QALYs are used routinely in the medical and public-health fields, whereas WTP is widely used in evaluating environmental and transportation-related risks.

Although they have been developed in different application areas, the QALY and WTP frameworks share important similarities: both are justified as representing the preferences of individuals, and both are summed across individuals to represent the social value of a change in health risk. However, the specific assumptions underlying the approaches differ in ways that produce systematic differences in the relative values of changes in risks. These differences may lead to different conclusions about whether a policy increases or decreases aggregate health risk.

This article provides an introduction to the QALY and WTP approaches, reviews the theoretical foundations underlying them, and examines the implications of differences in the foundations for ranking environmental and other health risks. In particular, it examines the effects of age, health, and income of the affected population on the value of reducing mortality risk.

The article is organized as follows. Section 2 reviews the theoretical assumptions of the two approaches. Section 3 examines the implications for valuing current mortality risk and aggregating values of mortality-risk changes across individuals. Section 4 describes empirical methods for estimating values...
under the two approaches. Section 5 discusses how health risks are aggregated across people or combined with other endpoints under the two approaches, and Section 6 concludes.

2. UTILITY-THEORETIC FOUNDATIONS

An individual experiences various “health states” over his or her lifetime. The time path of health states experienced, ending in death, is a “health profile.” Risks to health and/or longevity may be represented as lotteries (probability distributions) over alternative health profiles, and policies or other interventions that alter health risks alter the probabilities associated with experiencing different health profiles.\(^1\)

A utility function is any function that summarizes an individual’s preferences, in the sense that it assigns a higher number to a more preferred lottery. Both QALYs and WTP are justified as representing individual utility functions. QALYs assume that preferences over health and longevity depend only on health consequences, and do not depend on other characteristics of the individual or the risk.\(^2\) In contrast, WTP allows for the possibility that preferences over health outcomes depend on individual characteristics, such as wealth, as well as on characteristics of the risk, such as whether it is perceived to be uncontrollable, unfamiliar, or dreaded.

2.1. QALYs

The QALY framework provides a method for measuring the value of a health profile in terms of the duration of an equally preferred health profile free of any health impairment. The number of QALYs in a specified health profile is calculated as the quality-weighted lifespan:

\[
QALYs = \sum_{i=1}^{M} q_i T_i. \tag{1}
\]

In Equation (1), lifespan is divided into \(M\) periods that are indexed by \(i\). The periods are defined so that only one health state is experienced in each period. The duration of period \(i\) is \(T_i\) and the “health-related quality of life” (HRQL) associated with that period is characterized by a weight \(q_i\). The value of an intervention that affects health and/or longevity is measured as the difference in QALYs between the health profiles obtained with and without the intervention, as illustrated in Fig. 1.

The HRQL is a number that represents the quality of health.\(^3\) It is scaled so that a value of one

\(^1\) Note that a health profile experienced with certainty can be represented as a degenerate lottery that assigns probability one to the certain health profile and probability zero to all other profiles.

\(^2\) Technically, preferences over health quality and longevity must be “utility independent”\(^{(1)}\) of other characteristics of the individual and the risk.

\(^3\) Several terms, including health-related quality of life, health status, and functional status, are used in the literature to designate a variety of single and multidimensional measures of health.
corresponds to perfect or excellent health, and a value of zero corresponds to health that is equivalent to death (i.e., an individual would not care if he or she were to live the rest of his or her lifespan in such a state or die immediately). Typically, \( q \) is between one and zero, but values of \( q \) less than zero can be used to represent states of health that are worse than death.

The conditions under which QALYs represent a valid individual utility function were identified by Pliskin et al.\(^4\) These authors restrict their attention to the special case of chronic (constant) health states, for which Equation (1) simplifies to:

\[
\text{QALYs} = q \times T
\]  

(2)

where \( T \) is remaining lifespan and \( q \) is the HRQL for the constant health state in which the individual will live until death. In this case, QALYs represent a valid utility function for an individual if his or her preferences satisfy the following conditions.

1. **Mutual utility independence.** This condition has two parts: (a) preferences between lotteries on health states, holding duration of life constant, do not depend on remaining lifespan; and (b) preferences between lotteries on lifespan, holding health state constant, do not depend on health state. An example of part (a) is if an individual is indifferent between living 40 years in “good” health and a 70–30 lottery between living 40 years in “excellent” health or in “fair” health, she is also indifferent between living 25 years in “good” health and a 70–30 lottery between living 25 years in “excellent” or “fair” health.\(^4\) An example of part (b) is if an individual is indifferent between living 30 years and a 50–50 lottery between living 40 years and 25 years, with all years lived in “excellent” health, then she is also indifferent between living 30 years and a 50–50 lottery between living 40 years and 25 years, with all years lived in “fair” health. Mutual utility independence is necessary for utility to be represented as a product of separate health and longevity terms.\(^1\)

2. **Constant proportional tradeoff of longevity for health.** The fraction of remaining lifespan the individual would be willing to sacrifice to improve his or her health from one state to another does not depend on his or her remaining lifespan. For example, if an individual is indifferent between living 40 years in “fair” health and 30 years in “excellent” health, she is also indifferent between living 20 years in “fair” health and 15 years in “excellent” health. This condition implies that the HRQL associated with a health state does not depend on the length of time spent in that state.

3. **Risk neutrality over lifespan.** Holding health state constant, the individual prefers whichever lottery on longevity provides the greatest life expectancy. For example, the individual would prefer to live 41 years to a 50–50 lottery between living 50 and 30 years, and she would prefer that lottery to living 39 years (where all years are lived in the same health state, e.g., “excellent” health). A risk-adjusted form of QALY (which does not require risk neutrality) has also been developed\(^4\) but is rarely used in practice. In the risk-adjusted case, the simple and ethically appealing calculation of changes in social utility as the population sum of individual changes in QALYs is inconsistent with individual preferences. This follows because the value of a health profile to an individual is a nonlinear function of duration, and so the individual’s utility is not equal to the sum of his or her quality-weighted life years.

Recently, Bleichrodt et al.\(^5\) and Miyamoto et al.\(^6\) proposed alternative and simpler conditions that imply an individual’s preferences over lotteries on chronic health profiles can be represented by QALYs. One condition is that the individual is indifferent among all health states when his or her lifespan is zero (the “zero condition”). If, in addition, he or she is risk neutral over lifespan for each health state (which implies that longevity is utility independent of health state), then his or her preferences can be described by Equation (2).\(^5\) Alternatively, if his or her preferences for lotteries on lifespan holding health constant do not depend on the health state (i.e., lifespan is utility independent of health), then his or her preferences...
5 The form of risk-adjusted QALY consistent with the Pliskin et al.(4) assumptions is more restrictive than the form consistent with the Miyamoto et al.(6) assumptions. The Pliskin et al.(4) assumptions require that risk posture with respect to longevity, holding health constant, satisfy constant relative risk aversion (or constant relative risk proneness). The Miyamoto et al.(6) assumptions impose no constraint on risk posture with respect to longevity (one can be risk averse for some range of lifespan and risk seeking for another), and do not even require that an individual always prefers a longer lifespan.

6 The alternatives (c) and (d) can be obtained from alternatives (a) and (b) by changing the health state for the first five years from “good” to “excellent.”

7 In contrast, QALYs measure the value of a health profile relative to immediate death.
impose so little structure on preferences, HYEs have not been widely used in practice.

2.2. WTP

The WTP approach reflects conventional microeconomic principles. Anything over which an individual has preferences, including lotteries on health profiles, can be described as an “economic good.” An individual’s preference for one lottery over another can be represented in terms of a change in income or wealth, which can be used to purchase other goods.

There are two alternative measures of an individual’s willingness to trade money and health: WTP and willingness to accept (WTA). Consider the value to an individual with wealth $w_0$ of moving from health profile $H_0$ to a preferred health profile $H_1$. Her utility is a function of the health profile and wealth, $u(H, w)$. The value of the improvement may be measured as:

1. WTP for improvement (compensating variation), the value of $c_0$ satisfying $u(H_0, w_0) = u(H_1, w_0 - c_0)$. The name implies that the loss of wealth $c_0$ compensates for the gain in health, leaving the individual no better or worse off than without the health improvement.

2. WTA in place of improvement (equivalent variation), the value of $e_0$ satisfying $u(H_0, w_0 + e_0) = u(H_1, w_0)$. The payment is equivalent to the health gain, in that the individual is equally well off whether she obtains the payment or the health improvement.\(^8\)

Fig. 2 illustrates WTP and WTA for changes in current-period mortality risk, holding the lottery on health and survival in future time periods constant. The figure illustrates two indifference curves for the probability of surviving a specified period (e.g., the current year) and wealth available for spending on other goods. An indifference curve is defined as a set of points such that the individual judges all points along it to be equally desirable. Points above and to the right of the indifference curve are preferred, as they represent larger survival probability and/or greater wealth. Under plausible assumptions (described in Section 3.2), the indifference curves relating survival probability and income are downward sloping and convex, as illustrated.

The initial position with survival probability $p_0$ and wealth $w_0$ is labeled A. An increase in survival probability to $p_1$ would shift the individual to B, on a higher indifference curve. The individual’s WTP for

\(^8\)Note that WTA to forego an improvement from $H_0$ to $H_1$ is different than willingness to accept compensation for a reduction in health from $H_0$ to some less desired health profile. One can also define WTP to prevent a reduction from $H_1$ to $H_0$ and WTA to permit a reduction from $H_1$ to $H_0$ (see Fig. 2).
this increase in survival probability is given by the vertical distance between the two indifference curves at \( p_1, B - C \). Alternatively, the individual could achieve the same increase in utility by moving to point D, which involves no change in his or her survival probability but an increase in his or her wealth. The individual's WTA compensation in lieu of the survival improvement is given by the vertical distance between the two indifference curves at \( p_0, D - A \).

If the risk reduction \( p_1 - p_0 \) is small, the two indifference curves will be nearly parallel between \( p_0 \) and \( p_1 \) (indifference curves cannot intersect). In this case, WTP and WTA will be nearly identical. If the risk reduction is large relative to the curvature of the indifference curves, WTA and WTP may be substantially different, with WTA > WTP.\(^9\) For large changes in mortality risk, an individual's WTA compensation in place of an increase in survival probability may be much larger than his or her WTP for the same survival gain (note that WTP is limited by ability to pay, but WTA is not).

In principle, the choice of whether WTP or WTA is the appropriate measure of a change in risk may depend on the "property right" in the situation. If the individual having wealth \( w_0 \) is entitled only to the inferior health profile \( H_0 \), then it may be appropriate to compare his or her WTP for the improvement to \( H_1 \) with the costs of providing the improvement. Alternatively, if the individual is entitled to \( H_1 \), then it may be appropriate to compare his or her WTA to forego the improvement with the costs that can be saved by not providing \( H_1 \). At the social level, when the costs of reducing risk are born by the beneficiaries, this distinction breaks down. If starting at \( H_0 \), the question is whether the individuals' collective WTP for an improvement exceeds the cost of improvement and, if starting at \( H_1 \), whether their collective WTA compensation for an increase in risk is less than the costs saved by allowing the increase. The situation in which individuals are entitled to \( H_1 \) without paying for it is not logically available in this case.\(^{21}\)

In practice, separate estimates of WTP and WTA can be most easily obtained using contingent valuation or other approaches in which respondents are questioned about their choices in hypothetical situations (described in Section 4). In these cases, estimated WTA is often much larger than estimated WTP. Estimated WTA is often viewed as implausibly large, and so attention has focused on estimating WTP even when WTA might be conceptually more appropriate.\(^{21,22}\)

3. VALUING MORTALITY RISK

In many cases, the health effect of greatest concern is fatality. The effects of individual characteristics, including age, health, competing mortality risk, and income, on the value of reducing mortality risk differ systematically between QALY and WTP approaches. In this section, the effects of these characteristics on the value of reducing a specific current risk (defined as a probability of dying within the current period from a specified cause) are examined under each framework.

3.1. QALYs

The value of a change in a specific current mortality risk under the QALY approach is the change in the expected number of QALYs. It depends on life expectancy and expected future health state, but not (with limited exceptions described below) on income or other factors.

If the probability of dying from a specific cause in the current period is \( p \), the individual faces a lottery with a \( p \) chance of dying in the current period, and a complementary chance of surviving the specific risk and facing the lottery over health profiles that is determined by all the other health risks he or she faces in the current and future periods. Assuming the current period is one year, the health profile if the individual dies from the specific risk provides approximately one-half QALY (assuming he or she is equally likely to die at any time during the year and that his or her HRQL until then is nearly one). The value of a small reduction in the specific fatality risk is:

\[
V = \Delta p E(QALY) - \Delta p / 2
\]
where $\Delta p$ is the change in the specific risk and $E(QALY)$ is the expected number of QALYs if he or she survives the specific risk. Assuming the expected future QALYs are large compared with 1/2, the second term in Equation (3) can be neglected, yielding:

$$V \approx \Delta pE(QALY). \quad (4)$$

As shown by Equation (4), the value of reducing a specific mortality risk depends on the health lottery the individual faces if he or she survives that risk. Indeed, it is nearly proportional to the expected number of QALYs the individual will live if he or she survives. This implies that the value of reducing the specific mortality risk is directly related to the individual’s life expectancy conditional on surviving the specific risk, and to his or her expected future health state. For an individual who is likely to survive in very good health ($q \approx 1$), the value of reducing the specific mortality risk is proportional to life expectancy. For example, the conditional life expectancy of U.S. residents is about 58 years at age 20 and 18 years at age 65, and so the value of reducing a near-term mortality risk to a 20 year old is approximately 3.2 ($=58/18$) times as large as the value of a comparable risk reduction to a 65 year old. If future QALYs are discounted, the effect of life expectancy is attenuated. Using a recommended discount rate of 3% per annum, the relative value of reducing risk to a 20 year old would be about twice as large as the value of reducing risk to a 65 year old.

The effect of a competing mortality risk is to reduce the value of mitigating the specific mortality risk in direct proportion to the magnitude of the competing risk. This follows because the competing risk reduces the expected QALYs conditional on surviving the specific risk. For example, it has been suggested that the individuals who are at greatest risk of dying because of particulate air pollution face very large competing risks because of their age and cardiopulmonary impairments. The associated competing mortality risk may approach one per year, consistent with a life expectancy of less than one year. Under the QALY approach, the value of reducing the risk that such people will die from air pollution is relatively small because their life expectancy conditional on surviving the air pollution is small.

The value of reducing a specific mortality risk is also proportional to the individual’s expected future health. Hence, the QALY approach implies it is more valuable to reduce a current mortality risk for someone whose survival would be in very good health than for someone whose survival would be in impaired health. For example, the HRQL for life after a myocardial infarction has been estimated as about 0.9. Under the QALY approach, the value of reducing current mortality risk to someone who has survived a myocardial infarction is about 90% as large as the value of an identical risk reduction to someone who will survive with the same life expectancy but with no significant health impairment. Similarly, if people at risk of death from air pollution have low HRQL because of preexisting illness, the QALY value of reducing mortality risk from air pollution may be lower than the value of reducing risks to healthier people.

The relative value of reducing mortality risks to different individuals under the QALY approach is generally considered to be independent of individual economic circumstances, because life years (adjusted for health status) are counted equally regardless of personal characteristics. However, this claim must be qualified, as wealth can have several effects on HRQL, which represents the rate of substitution between longevity and health quality.

First, HRQL may depend on individual characteristics and circumstances. For example, the utility consequence of a health impairment may depend on the individual’s ability to mitigate it, which may depend on economic circumstances. If the effects of an adverse health condition on individual well-being can be substantially offset using market goods (e.g., personal assistants or mechanical devices), an individual’s well-being in that state may be positively related to his or her wealth or income. However, since HRQL measures utility in the impaired health state relative to utility in perfect health, the effect of wealth on HRQL will depend on the relative degree to which it improves well-being in the two states. Under the assumption that QALYs are a utility function for health and longevity, the incremental effect of wealth on welfare is positively associated with health and longevity, except in the implausible case in which incremental wealth is more valuable as a bequest than in life. Limited empirical evidence also suggests that the marginal utility of wealth is smaller in impaired health states than in full health.

---

10 Life expectancy of people who die from air pollution is estimated to be on the order of months to years, although for those with chronic obstructive pulmonary disease, the mortality displacement may be on the order of weeks to months.

11 Sloan et al estimate that having multiple sclerosis (MS) reduces the marginal utility of income by a factor of 0.67 (estimated
Second, under the approach recommended by an expert panel,(2) the effects of health status on earnings capability and income are incorporated in HRQL. The effect of a health impairment on income is likely to depend on both income and the individual’s job. Individuals whose income is more sensitive to health status may have a smaller HRQL for the same health impairment (e.g., a physical disability might cause a greater income loss to a construction worker than to a writer).

For evaluating the social value of changes in health risk, the effects of income or other individual characteristics on HRQL can be eliminated by valuing all changes using population-average values of HRQL. Indeed, this is the recommended practice.(2) However, if HRQL depends on income, this approach does not aggregate individual changes in welfare and so may lead to ranking health interventions in an order different than the affected individuals would rank them. Note that the same approach—using population-average values—can be (and usually is) used to remove the effect of income differences on WTP.

Since QALYs depend only on the duration and severity of health effects, the value of a risk reduction is independent of other characteristics of the risk, such as whether it is perceived as controllable or dreaded. In principle, the HRQL associated with a health state might be allowed to depend on these characteristics, but this extension has not been investigated.

3.2. WTP

Under the WTP approach, the value of reducing mortality risk is measured as the “value of a statistical life” (VSL). VSL is an individual-specific value defined as the marginal rate of substitution between mortality risk and wealth or income, that is, the individual’s WTP for a small reduction in mortality risk divided by the risk change, which is equivalent to the WTA for a small increase in mortality risk divided by the risk change and to the slope of the indifference curve illustrated in Fig. 2 at the individual’s wealth and risk level.

VSL depends on wealth, current mortality risk, and the lottery over future health profiles and individual faces. The standard model(31–33) assumes that the individual maximizes his or her expected utility:

\[ EU(p, w) = (1 - p) u_a(w) + pu_d(w) \tag{5} \]

where \( p \) is the individual’s chance of dying during the current period and \( u_a(w) \) and \( u_d(w) \) represent his or her utility as a function of wealth conditional on surviving and not surviving the period, respectively. The function \( u_d(\bullet) \) incorporates the individual’s preferences for bequests and can incorporate any financial consequences of dying (such as medical bills or life insurance benefits). In this single-period model, wealth and income are treated as equivalent. In multi-period models, the difference between wealth and income and the opportunities for future earnings can be important.

The individual’s VSL is derived by differentiating Equation (5), holding utility constant, to obtain:

\[ VSL = \left. \frac{dw}{dp} \right|_{EU=kw} = \frac{u_a(w) - u_d(w)}{(1 - p)u'_a(w) + pu'_d(w)} = \frac{\Delta U}{EU} \tag{6} \]

where prime indicates first derivative.

The numerator in Equation (6) is the difference in utility between surviving and dying in the current period. The denominator is the expected marginal utility of wealth, that is, the incremental utility associated with additional wealth conditional on surviving and dying in the current period, weighted by the respective probabilities. Assuming that survival is preferred to death (i.e., \( u_a(w) > u_d(w) \)) and that greater wealth is preferred to less (i.e., \( u'_a(w) > 0, u'_d(w) \geq 0 \)), both numerator and denominator are positive and so VSL is positive and the indifference curves in Fig. 2 slope downward.

Under the WTP approach, as under the QALY approach, the value of reducing a specific mortality risk in the current period depends on life expectancy, competing mortality risk, and the individual’s health if he or she survives the specific risk. In addition, the value under the WTP approach also depends on baseline risk and on income or wealth.

For example, consider the effect of baseline (total) risk on VSL. It is natural to assume that \( u'_a(w) > u'_d(w) \), that is, the increased utility provided by greater wealth is larger if the individual survives and has the opportunity to spend it. If so, an increase in the baseline risk \( p \) decreases the expected utility cost of spending (the denominator in Equation (6)). The utility associated
with survival (the numerator in Equation (6)) is unaffected by baseline risk, so the individual would be willing to spend more to reduce his or her mortality risk. For small changes in risk, this “dead-anyway” effect is small. Assuming that \( u'_1 \geq 0 \) (i.e., the individual prefers more wealth to less, even if he or she dies), the proportional effect on VSL of a change in baseline risk is less than the proportional change in the survival probability \((1 – p)\).

The value of reducing the specific mortality risk is smaller if the individual also faces a competing mortality risk. The existence of a competing mortality risk reduces the magnitudes of the numerator and the denominator in Equation (6). The numerator decreases because the total probability of survival is smaller, and the denominator decreases because of the dead-anyway effect. It can be shown, however, that the effect in the numerator dominates, and so competing mortality risk reduces WTP to reduce the specific mortality risk.\(^{35}\)

VSL may depend on the individual’s future health if he or she survives the specific mortality risk, but the sign of the effect is ambiguous. Survival in good health rather than poor increases the value of the numerator in Equation (6). However, if the marginal utility of wealth is higher in good health than in poor health, the value of the denominator is larger and the effect on the ratio is indeterminant. As noted above, limited empirical evidence suggests that the marginal utility of income is smaller in a state of chronic health impairment,\(^{28,29}\) and some empirical studies suggest that VSL is larger for people with cancer,\(^{36,37}\) or angina,\(^{37}\) than for people without those impairments.

As with most goods, WTP for reduction in mortality risk depends on ability to pay and is likely to increase with wealth. The assumption that additional wealth is more valuable in life than as a bequest (i.e., \( u'_w(w) > u'_d(w) \)) implies that the numerator of Equation (6) increases with wealth. Individuals are generally averse to financial risk. If so, the denominator declines with wealth (the second derivatives of \( u_w(w) \) and \( u_d(w) \) are negative), and VSL increases. If the individual is indifferent to financial risk, the denominator is constant and again VSL increases with wealth. Only in the implausible case in which the individual prefers to bear greater financial risk (for the same expected return) can the denominator increase with wealth, making the effect on VSL indeterminate.\(^{13}\)

The effect of life expectancy on VSL is influenced by two competing factors. A greater life expectancy increases the utility of surviving the current period (the numerator in Equation (6)). Greater life expectancy may also increase the denominator, because of the desire to save wealth for consumption in future periods, or because the opportunity cost of current spending to reduce mortality risk is larger for individuals with a longer investment horizon. The effect of life expectancy also depends on whether the individual is able to borrow against future income and on any difference between the rate at which he or she discounts future utility and the rate of return to savings.

A number of investigators have developed theoretical models to examine how VSL varies over an individual’s life cycle. These models extend the one-period model described in Equation (5) by assuming the individual seeks to maximize the expected discounted value of the utility of consumption:

\[
EU = \sum_{t=0}^{\infty} p_t \delta^t u(c_t) \tag{7}
\]

where \( p_t \) is the probability of surviving at least to age \( t \), \( c_t \) is consumption at age \( t \), and \( \delta \) is the individual’s discount factor (i.e., \( \delta = 1/(1 + r) \) where \( r \) is the rate at which the individual discounts future utility).

In models that assume an individual can borrow against future earnings, VSL declines monotonically with age. Under this assumption, Shepard and Zeckhauser\(^{38}\) calculate that VSL for a typical American worker falls by a factor of three from age 25 to age 75. If individuals can save but not borrow, VSL peaks near age 40 and is less than half as large at ages 20 and 65.

Ng\(^{39}\) suggests that individuals may discount their future utility at a rate smaller than the rate of return to financial assets, whereas Shepard and Zeckhauser\(^{38}\) assume these rates are equal. If the utility-discount rate is less than the rate of return, individuals should save more when they are young and consume more when old. Under these conditions, VSL may not peak until age 60 or so.\(^{39}\) Even if they discount future utility at the rate of return, prudent\(^{40}\) individuals might be anticipated to save more and spend less on reducing mortality risk when they are young because of the greater range of financial contingencies they face.

WTP may depend on characteristics of the risk other than the probabilities and possible health

---

13 Positive effects of baseline risk and wealth on VSL are sufficient conditions for the convexity of the indifference curves in Fig. 2.
outcomes. Limited empirical evidence suggests that average WTP to reduce fatality risks may be somewhat larger for risks that are perceived as involuntary, uncontrollable, unfamiliar, or dreaded.\(^{41-44}\)

For evaluating social programs, it is possible to ignore the effects of individual differences in wealth or other factors that are considered ethically inappropriate by replacing individual VSLs with a value that is obtained by averaging over the objectionable characteristics. This approach is often taken in practice, where differences in wealth and health quality are generally ignored. An alternative approach is to consider how individuals might choose to incorporate differences in wealth and other factors in allocation of social resources if they were to make the decision behind a Rawlsian veil of ignorance before they knew their own characteristics. Pratt and Zeckhauser\(^{34}\) use this approach to argue that the appropriate VSL for use in social policy choices increases with income, although at a smaller rate than empirical estimates suggest. They also argue that differences in VSL due to differences in baseline risk (the deadAnyway effect) should not be incorporated.

4. METHODS FOR ESTIMATING VALUES

Information about preferences, in the form of HRQL or WTP, can in principle be obtained using stated-preference or revealed-preference methods. Stated-preference methods have been used to estimate both HRQL and WTP, but to date revealed-preference methods seem to have been used only to estimate WTP.

Stated-preference methods rely on asking individuals either to report their preferences directly or to report how they would behave in a specified hypothetical situation. They are extremely flexible, as individuals can be questioned about how they would choose in a great variety of hypothetical situations. The hypothetical nature of the choice is also the greatest weakness of these methods, as individuals may be unfamiliar with the choices and have inadequate incentive or opportunity to provide thoughtful answers.

Revealed-preference methods rely on observing behavior in situations that are more consequential than answering survey questions. They assume that people act in their own best interest and thus the chosen alternative must be preferred to the rejected alternatives. In revealed-preference studies, subjects have an incentive, and may have the opportunity, to seek information about the alternatives and to consider the choice carefully. Nevertheless, individuals may be poorly informed about the differential health risks associated with the choices they face. Also, although the analyst observes the alternatives that individuals choose, he or she does not observe the alternatives they reject and the attributes of those alternatives.

This section provides a brief overview of the methods used to estimate HRQL and WTP. The identification of possible health states and probability distribution over time spent in each state that is required for calculating QALYs can be developed using risk assessment methods.

4.1. QALYs

HRQL is typically elicited directly or calculated from a generic health utility scale. The generic scales are themselves calibrated using direct elicitation.

4.1.1. Direct Elicitation

The HRQL may be elicited from individuals directly, using any of several question formats: standard gamble, time tradeoff, visual analog scale, and person tradeoff. In general, HRQL for a health state is elicited assuming the health state will be chronic (constant).

The standard gamble (SG) format requires the respondent to indicate the smallest chance of survival in perfect health he or she would accept in a lottery where the alternative outcome is immediate death. This may be motivated by considering a surgery that would alleviate a health impairment without affecting longevity, except for the chance of dying in surgery. For example, if the respondent is indifferent between living 20 more years in a particular impaired health state and a lottery that offers her a 75% chance of living 20 more years in perfect health and a complementary chance of immediate death, the value of q for the impaired health state is 3/4, and both the certain health profile and the lottery offer an expected value of 15 QALYs.

The time tradeoff (TTO) format requires the respondent to indicate the number of years in perfect health (with q = 1) he or she considers to be indifferent to a specified chronic health profile. For example, if the respondent indicates that she is indifferent between living 20 years in a particular impaired health state and 15 years in perfect health, the value of q for the impaired health state is calculated as 15/20 = 3/4. Both health profiles offer 15 QALYs.

The visual analog scale (VAS) is a linear scale with one end representing perfect health and the other
representing health states as bad as death. The respondent is asked to place a mark on the scale representing how desirable the specified health state is to him or her, relative to the endpoints. A similar verbal format may be used where the respondent is asked to report a number representing his or her preference for the health state between 0 and 100, where 0 represents a state as bad as death and 100 represents perfect health.

The person tradeoff (PTO) format asks the respondent to consider the relative value of improving health for people in different health states. For example, she might be asked to judge the relative value of extending longevity for people in different health states, e.g., if one were to choose between extending the life of 1,000 healthy people for a year and extending the life of x blind people for a year, for what value of x would she be indifferent? The HRQL of living with blindness is estimated as $1,000/x$. Alternatively, the respondent might be asked to judge the relative value of improving health for people in one state and extending life for people in another state, for example, if one were to choose between extending the life of 1,000 healthy people for a year and restoring the site of z blind people for a year, for what value of z would she be indifferent? In this case, the HRQL of living with blindness is estimated as $1 - (1,000/z)$.

Risk-tradeoff questions have been used to evaluate preferences for environmental and motor-vehicle related risks. In a risk-tradeoff question, respondents are asked to choose between situations offering higher risks of one health outcome (e.g., chronic bronchitis) and lower risks of another (e.g., motor-vehicle fatality). The risk-tradeoff approach is similar to SG. A respondent who is indifferent between reducing his motor-vehicle-fatality risk by 3 per 10,000 and his risk of chronic bronchitis by 1 per 1,000 can be interpreted as having an HRQL for chronic bronchitis of 0.7.

If the conditions under which QALYs provide a valid utility function are satisfied, and an individual’s answers to elicitation questions are consistent with his or her utility function, then both SG and TTO formats should yield exactly the same value. The claim that SG incorporates risk preferences whereas TTO only captures preferences about risk-free outcomes is incorrect. In practice, the results of SG and TTO elicitations differ, perhaps because individuals’ preferences are not exactly consistent with the required conditions and because the formats make different aspects of the health profiles more salient: SG emphasizes risk and uncertainty, while TTO emphasizes relative preferences for near-term and future health. In practice, SG values may be slightly larger than TTO values. VAS values, because they are not tied to an explicit decision, have a weaker utility-theoretic justification. In practice, however, they may be more reliably assessed (i.e., vary less on repeat measurement) than TTO or SG values. VAS values are typically smaller than TTO or SG values, but are sometimes adjusted using an empirically estimated formula to approximate the results of TTO or SG formats. Unlike the other methods, PTO has the potential to incorporate judgments about distributional equity. PTO measures preferences over other people’s health, and so values elicited using PTO need not correspond to values of HRQL that represent an individual’s preferences for his or her own health.

An important question in eliciting HRQL is whose values to elicit? Possible respondents include those randomly sampled from the general public, individuals experiencing the health states of interest, and health-care providers or others knowledgeable about the health state. Experience suggests that individuals in an impaired health state assign a larger HRQL to that state than do healthier individuals. Whether this reflects improved understanding of the condition by people experiencing it or adaptation to adverse circumstances is not clear. All the choice-based elicitation methods require comparing two health states, at least one of which is not currently experienced by the respondent at the time of elicitation.

4.1.2. Generic Health Utility Scales

A number of generic health utility scales have been developed. These scales can be used to describe health states in terms of their levels on several attributes, and the HRQL associated with the state may be obtained from a table or calculated using an arithmetic formula. In principle, all such scales are examples of multi-attribute utility functions, although the extent to which the scales are explicitly based on multi-attribute utility theory varies. The scales have been calibrated by fitting them to preference values elicited using one or more of the direct methods reviewed above. Among the more popular generic health utility scales are the Health Utilities Index, the EuroQol EQ-5D, and the Quality of Well-Being Index.

The Health Utilities Index Mark 3 (HUI3) classifies health states using a system of eight attributes:

---

14 Risk-adjusted QALYs may be written in a form where the answers to SG and TTO questions are not equal but are related to each other by a known transformation.
The Quality of Well-Being Index (QWB)\(^{(52)}\) is one of the earliest of the generic health utility scales. It describes health using three attributes—mobility, physical activity, and social activity—plus an attribute consisting of descriptions of symptoms or “problem complexes.” Like the EQ-5D, the HRQL is an additive function of the attribute levels. The attribute levels and values are reported in Table III. In addition, there are 27 different symptoms or complexes with values ranging from 0 (no symptoms) to −0.727 (death). An individual with mobility Level 2 (in hospital, health related), physical activity Level 4 (no limitations for health reasons), social activity Level 3 (limited in major (primary) role activity, health related) with symptom and problem complex 10 (general tiredness, weakness, or weight loss, value = −0.259) would have HRQL = 1 − 0.090 − 0.000 − 0.061 − 0.259 = 0.590.

### Table I. Health Utilities Index Mark 3

<table>
<thead>
<tr>
<th>Level</th>
<th>Vision</th>
<th>Hearing</th>
<th>Speech</th>
<th>Ambulation</th>
<th>Dexterity</th>
<th>Emotion</th>
<th>Cognition</th>
<th>Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>0.98</td>
<td>0.95</td>
<td>0.94</td>
<td>0.93</td>
<td>0.95</td>
<td>0.95</td>
<td>0.92</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>0.89</td>
<td>0.89</td>
<td>0.89</td>
<td>0.86</td>
<td>0.88</td>
<td>0.85</td>
<td>0.95</td>
<td>0.90</td>
</tr>
<tr>
<td>4</td>
<td>0.84</td>
<td>0.80</td>
<td>0.81</td>
<td>0.73</td>
<td>0.76</td>
<td>0.64</td>
<td>0.83</td>
<td>0.77</td>
</tr>
<tr>
<td>5</td>
<td>0.75</td>
<td>0.74</td>
<td>0.68</td>
<td>0.65</td>
<td>0.65</td>
<td>0.46</td>
<td>0.60</td>
<td>0.55</td>
</tr>
<tr>
<td>6</td>
<td>0.61</td>
<td>0.61</td>
<td>na</td>
<td>0.58</td>
<td>0.56</td>
<td>na</td>
<td>0.42</td>
<td>na</td>
</tr>
</tbody>
</table>

HRQL = 1.371 \(b_1 \ast b_2 \ast b_3 \ast b_4 \ast b_5 \ast b_6 \ast b_7 \ast b_8\) − 0.371

Note: na = not applicable (attribute has only five levels).

Source: Reference 53.

The HUI formula and attribute levels are reported in Table I. As an example, an individual with functioning at the highest level on all attributes except vision (Level 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) and ambulation (Level 4: able to walk only short distances with walking equipment, and requires a wheelchair to get around the neighborhood) would have HRQL = 1.371 \((0.89 \ast 1 \ast 1 \ast 0.73\ast 1 \ast 1 \ast 1)\) − 0.371 = 0.52.

The EuroQol EQ-5D classifies health states using a system of five attributes: mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Each attribute has three levels, yielding a total of 243 health states. Two additional attributes, dead and unconscious, have been added for a total of 245. Values for EQ-5D states have been elicited using TTO and VAS format questions in numerous European and Nordic populations. As an approximation, the HRQL can be represented as an additive function of the attribute levels, as shown in Table II. An individual with impaired mobility (Level 2: some problems in walking about) and some pain/discomfort (Level 2: moderate pain or discomfort), with all other attributes at their highest levels, would have HRQL = 1 − 0.069 − 0.123 − 0.081 = 0.73. The five-attribute classification system is supplemented with a single holistic question in which the respondent is asked to rate his or her current health on a visual analog scale.\(^{(54)}\)

### Table II. EuroQoL EQ-5D

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>−0.0</td>
<td>−0.069</td>
<td>−0.314</td>
</tr>
<tr>
<td>Self-care</td>
<td>−0.0</td>
<td>−0.104</td>
<td>−0.214</td>
</tr>
<tr>
<td>Usual activity</td>
<td>−0.0</td>
<td>−0.036</td>
<td>−0.094</td>
</tr>
<tr>
<td>Pain/discomfort</td>
<td>−0.0</td>
<td>−0.123</td>
<td>−0.386</td>
</tr>
<tr>
<td>Anxiety/depression</td>
<td>−0.0</td>
<td>−0.071</td>
<td>−0.236</td>
</tr>
</tbody>
</table>

HRQL = 1 − sum of relevant item weights. The additional constant 0.081 is subtracted if any attribute is Level 2, and 0.269 is subtracted if any attribute is Level 3.

Source: Reference 55.
the change in probability, and in some cases WTP is not even statistically significantly related to the magnitude of risk reduction.\(^{57}\)

In a recent study, Corso \textit{et al.}\(^{61}\) investigated the extent to which difficulties in communicating small risks to survey respondents might account for the inadequate sensitivity of estimated WTP to magnitude of the risk reduction. In separate subsamples, they elicited WTP for a side-impact automobile airbag that was described as reducing the annual chance of dying in an automobile crash by either 5/100,000 or 10/100,000. Respondents were presented with one of three visual aids (a logarithmic risk ladder, a linear risk ladder, or a field of 25,000 dots), or with no visual aid. Corso \textit{et al.}\(^{61}\) found that estimated WTP was proportional to the stated risk reduction for the respondents who were presented with the 25,000 dots, and WTP was close to proportional for the respondents presented with the logarithmic risk ladder. In contrast, WTP was not significantly related to the stated magnitude of the risk reduction for respondents who were not provided with a visual aid.

### 4.2. WTP

WTP for reductions in health risk can be estimated using either stated- or revealed-preference methods. These are described in turn.

#### 4.2.1. Stated Preferences

The most commonly used stated-preference method is contingent valuation (CV), in which survey respondents are asked to choose between alternatives that differ in the attribute to be valued and in cost. CV has been used to value a range of health risks, beginning with Acton’s study of emergency-response services for heart attacks.\(^{56}\) The most common application has been to transportation risks, although risks associated with food, medical technologies, hazardous waste, and other sources have also been examined.\(^{57}\)

CV results can be sensitive to apparently inessential aspects of the choice (e.g., question ordering and the format in which risks are presented) but insensitive to essential aspects, such as the quantity of the good to be valued.\(^{57-61}\) CV estimates of WTA are often much larger than estimates of WTP, perhaps because of framing effects.\(^{21,22}\)

For estimating VSL, the apparent insensitivity of WTP to the magnitude of risk reduction is important because the estimated VSL (WTP divided by risk reduction) will strongly depend on the magnitude of the reduction specified in the survey. Standard theory (Section 3.2) suggests that WTP for mortality-risk reduction should be nearly proportionate to the magnitude of the change in probability (i.e., VSL should be insensitive to small changes in baseline risk and wealth). The modest number of studies that have tested for sensitivity to magnitude have found that estimated WTP is usually less than proportionate to the reduction speciﬁed in the survey. Standard theory

#### 4.2.2. Revealed Preferences

Revealed-preference studies of WTP for health risk reductions require that individuals choose between alternatives that differ in health risk and monetary consequences. Most have examined the incremental pay workers receive for accepting hazardous jobs. Choices among consumer products (e.g., cigarettes, smoke detectors, automobiles) and the use of protective equipment (e.g., seatbelts, motorcycle helmets) have also been examined.\(^{62}\)

Studies of compensating wage differentials suffer from data and statistical limitations. Fatality risk is usually based on industry or occupational averages, which are likely to conceal much variation between jobs\(^{63}\) and workers.\(^{64}\) Inability to control for all the other job and worker attributes that may be correlated with fatality risk leads to potential biases.\(^{65}\) For example, many studies do not control for nonfatal-injury risk. This omission is likely to bias estimated VSL upward because part of the observed wage differential compensates for injury risk, which is positively correlated with fatality risk. The bias is estimated as 20–150% using actuarial data on risks to U.S. workers\(^{66}\) and 100% using survey data on perceived risks to Taiwanese workers,\(^{67}\) although a recent meta-analysis\(^{68}\) suggests the bias is negligible.

Although it may appear that studies of compensating wage differentials estimate workers’ WTA

### Table III. Quality of Well-Being Index

<table>
<thead>
<tr>
<th>Step</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mobility</td>
</tr>
<tr>
<td>5</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>-0.062</td>
</tr>
<tr>
<td>3</td>
<td>na</td>
</tr>
<tr>
<td>2</td>
<td>-0.090</td>
</tr>
<tr>
<td>1</td>
<td>na</td>
</tr>
</tbody>
</table>

\(HRQL = 1 - \text{sum of scale weights} - \text{additional term for relevant symptom and problem complex.}\)

\(\text{Note: na = not applicable (not all scales use all steps).}\)

\(\text{Source: Reference 55.}\)
compensation for job risk, these studies equally measure workers’ WTP to reduce risk. A worker is assumed to prefer the job he or she holds to all the alternative jobs available to him or her. Implicitly, he or she reveals that his or her WTA compensation to bear additional risk is larger than the incremental pay offered by more dangerous jobs, and that his or her WTP for risk reduction is smaller than the pay cut he or she would take by choosing a safer job.

5. AGGREGATION OF HEALTH RISKS

A fundamental difficulty in defining the social value of changes in health risk is the absence of a clear criterion for weighing gains and losses to different people. Standard microeconomic theory assumes it is impossible to measure utility or to compare utility gains between people. From this perspective, the choice between using years of healthy life and monetary units as a standard for comparing utility changes between individuals is arbitrary.

Choosing a standard for comparing utility changes among individuals is analogous to choosing which of a possibly infinite set of Pareto efficient allocations of welfare is the most socially preferred. The choice may be based on which standard is judged to be more equitable, but other methods of addressing equity implications are also available. For example, whether utility is measured in QALYs or monetary units, changes in utility to different individuals can be weighted by some function of individual characteristics; for example, QALYs gained by people with poor health could be weighted more heavily than gains to people in good health, and WTP of people with low income could be given greater weight than WTP of people with high income. Alternatively, interpersonal differences in QALYs or WTP that are judged to result from ethically inappropriate factors can be removed by ignoring variation in these factors and using population-average values of QALYs or WTP. Ignoring the effects of individual characteristics comes at the cost of weakening the utility-theoretic basis for the measure and may lead to evaluations that contradict individual preferences.

The difference in unit also affects the ease with which changes in health risk can be compared with the resource costs of a policy and effects on other attributes of concern, such as environmental quality. WTP values can be easily combined with monetary measures of the costs of a policy and WTP for changes in other attributes, enabling one to identify an option that maximizes net benefits (as measured in monetary units). In principle, one could also measure the value of resource costs, changes in environmental quality, and other attributes in QALYs, allowing identification of the policy that maximizes net benefits (as measured in QALYs). In practice, however, values of other attributes have not been measured in QALYs, and QALYs have been restricted to use in cost-effectiveness analysis, in which policies are evaluated by the cost per QALY produced. Cost-effectiveness analysis can be used to identify the least costly methods to provide health, but it cannot be used to determine whether the health gains from a particular policy justify the costs.

If a constant WTP per QALY ratio existed, it might be possible to translate from one approach to the other, much as one can translate monetary values between currencies using established exchange rates. However, the qualitatively different effects of life expectancy, health state, baseline risk, and wealth on the QALY and WTP measures of the value of current fatality risk imply that individuals cannot be expected to have a constant rate of substitution between QALYs and wealth. For example, even though life expectancy and future QALYs decline with age (after infancy), there appears to be a substantial age range over which VSL is constant or increasing. Over this range, the individual’s WTP per QALY must increase with age.

6. CONCLUSIONS

Even though QALYs and WTP are both justified as representing individual preferences, these two prominent methods for quantifying the social value of changes in health risk differ in their theoretical foundations, the unit by which health is measured, and in the relative values they assign to different health risks.

The different assumptions underlying QALYs and WTP have systematic effects on the quantified value of changes in current mortality risk. These are summarized in Table IV. Under both approaches, the value of reducing a specific current mortality risk is smaller when the individual faces a competing

---

15 An allocation is Pareto efficient if it is impossible to improve someone’s welfare without reducing someone else’s welfare. In dividing a cake among individuals, each of whom has an unlimited appetite and no concern for others, every allocation in which all the cake is eaten is Pareto efficient.

16 Similarly, individuals cannot be expected to have a constant value per discounted statistical life year.
mortality risk. Greater life expectancy and the absence of comorbidities increase the value of reducing a current mortality risk under the QALY approach, but the effects of these factors are theoretically ambiguous under the WTP approach. In contrast, baseline risk increases the value of reducing the current mortality risk under the WTP approach, but has no effect under the QALY approach. WTP depends on wealth, but QALYs are generally considered to be independent of wealth, subject to some qualifications regarding the extent to which wealth may help in coping with health impairment and the effect of health impairments on income.

QALYs are based on the assumption that preferences over health risks depend only on the probabilities of each health outcome. Preferences over health risks associated with other aspects of the risk, such as controllability and dread, cannot easily be incorporated in a QALY measure. In principle, the WTP approach can easily incorporate such preferences, although there have been few empirical attempts to estimate the effects of these factors.

Methods for estimating QALYs and WTP have been developed and applied to a wide set of health risks. Estimates of QALYs are likely to be less variable across people and studies than estimates of WTP because the QALY framework imposes greater constraints. Estimates of the duration of health states are typically based on modeling and the estimated HRQL for each health state may be partially constrained by comparison with estimates of HRQL for other health states that can be judged as better or worse. Because HRQL is scaled relative to perfect health and death, standard-gamble estimates of HRQL are typically obtained using probabilities on the order of 1 in 100 or larger, which are likely to be more comprehensible to survey respondents than the probabilities on the order of 1 in 1,000 or less that are often required for estimates of WTP. Small-sample studies that have elicited both WTP and HRQL using stated-preference methods suggest that WTP is less reliable.\(^{71,72}\)

Fundamentally, the choice between using QALY and WTP approaches depends on judgments about what constraints on individual preferences should be imposed and what factors should be considered in forming social judgments. QALYs impose substantial and somewhat unrealistic constraints on the form of individual preferences and combine preferences across people on a relatively egalitarian basis. In contrast, WTP imposes few constraints on individual preferences and gives relatively greater weight to more affluent sectors of society.

**ACKNOWLEDGMENTS**

An earlier version of this article was presented at the December 2000 meeting of the Society for Risk Analysis. I thank Patrick Hofstetter and three anonymous reviewers for helpful comments, and the U.S. Department of Agriculture Economic Research Service (43-3AEM-9-80125) for financial support.

**REFERENCES**


