Summary

The Value of Life – The Rise and Fall of a Scientific Research Programme

TØI Report 1531/2016
Author: Rune Elvik
Oslo 2016 239 pages English language

Research designed to obtain a monetary valuation of life and limb has long traditions. The value of saving a life was originally estimated as the capitalised value of output lost as a result of a premature death. This approach was abandoned around 1970 and a new theoretical foundation for the monetary valuation of life and limb was proposed: the willingness-to-pay (WTP) approach. Since that time, a large number of studies of willingness-to-pay for reduced risk of death have been made. This report reconstructs the history of this research by applying the methodology of scientific research programmes, developed by Imre Lakatos. The methodology of scientific research programmes can explain why a field of research can continue to exist despite that fact that many of its findings are difficult to make sense of and diverge enormously. Monetary valuation of life and limb fits this description. Estimates of the value of preventing a death, often referred to as the value of a statistical life, vary enormously. Some of this variation can be explained according to economic theory, but quite a lot cannot. One response to this state of affairs has been a reformulation of relevant parts of economic theory which allows for re-interpreting findings that were initially regarded as anomalous, but become theoretically plausible when re-interpreted. The reformulation of theory has almost gone to the point of making any finding theoretically plausible. It seems clear the valuation research based on willingness-to-pay originally had a goal of finding a single, uniform vale of a statistical life. Today, this objective has been given up and valuation research appears to live happily with the enormous diversity in estimates of the value of a statistical life.

This report is the final report of the project: “A historical reconstruction of research on the monetary valuation of transport safety by means of Imre Lakatos’ methodology of scientific research programmes”. The project has been funded by the Research Council of Norway as part of the TRANSIKK research programme.

Starting point: The values are all over the place

The starting point for the study is the observation that estimates of the value of a statistical life are all over the place. The value of a statistical life is the monetary value of a reduction in the risk of death, which statistically corresponds to the prevention of one death. Many studies have been made to estimate the value of a statistical life. The results of these studies vary enormously, from less than 5,000 US dollars to close to 200 million US dollars.

One might think that a field of research producing such diverse estimates, all of which are intended to measure the same thing, would be abandoned. However, valuation research continues to prosper and new studies are published quite frequently. This forms the background of the first main research problem of this study:

How can one explain that a field of research, producing enormously diverse results, some of which appear to contradict the theoretical foundations of the research, continues almost as if the contradictory results did not exist?

There exists a theory of science which seems well-suited to explaining the apparent paradox of valuation research. That is the methodology of scientific research programmes, developed by philosopher Imre Lakatos. Hence, the second main research problem is:

Can the methodology of scientific research programmes help to better understand, and possibly explain, the historic development of research on the monetary valuation of life and limb, in particular valuation of improving road safety?

Before summarising the answers to these questions, the context of the study will be briefly explained.

Do we need a monetary valuation of life and limb?

Many people find the very idea of assigning a monetary value to life or health objectionable, or at least strange. Many will ask: Do we really need these monetary valuations and what are their principal uses?

In this report, it is regarded as a basic and self-evident fact that trade-offs between different goods and objectives are made, and have to be made. Whenever an individual or government makes a decision about how much to spend on activities or measures that reduce the risk of death or injury, a trade-off is made between this good and other goods. It is impossible to avoid making trade-offs, simply because the resources at our disposal are limited. Therefore, the idea that life has an infinite value makes no sense.

However, it does not follow from this that a monetary value must by necessity be assigned to human life and health. It is entirely possible to avoid doing so and still make intelligent trade-offs between human life and health and other goods. Priorities between measures intended to reduce mortality or improve health can be set according to cost-effectiveness. The less a measure costs per fatality prevented, the higher should be its priority. The applicability of cost-effectiveness analysis is highly limited. Cost-effectiveness is undefined if a measure has an effect both on fatalities and injuries. Moreover, cost-effectiveness does not tell us when a measure is too expensive.

By explicitly assigning a monetary value to life and health, one may in principle:

1. Make trade-offs between safety and other policy objectives, like mobility or environmental protection – provided these objectives are also stated in monetary terms.
2. Determine if a measure improves societal welfare, which means that its benefits are greater than the costs, so that compensation preventing anyone from a net loss is in principle possible. It is assumed that the monetary valuation of non-market goods reflects their impacts on welfare.
3. Find the optimal use of policy instruments, for example road safety measures. The use of a set of measures is optimal when net benefits are maximised. In principle, this guideline can also be used to determine the size of a budget (it should be exactly large enough to cover the costs of all measures whose marginal benefits are equal to or greater than marginal costs).

These points indicate ways in which policy making can be informed by monetary valuation, but not without monetary valuation.

The definition, measurement and valuation of risk

Monetary valuation of life and limb refers to the valuation of changes in the risk of dying or of sustaining an injury. Most valuation studies state the risk of dying as a population mean fatality rate for a specific cause of death, for example that the current mortality rate in road accidents in Norway is close to 3 per 100,000 inhabitants per year.

This is the current population average for Norway. Risk varies in the population, by gender, by age groups, by place of residence and according to how much one travels by road. A population average may therefore be misleading as an estimate of the risk referring to a certain group of the population or a specific individual. Some valuation studies have therefore asked respondents about their subjective estimates of risk, i.e. what they think their own risk is. The majority of valuation studies has relied on statistical estimates of risk based on official statistics, applying either to the entire population of a country or a group of the population.

The good being valued in a valuation study is a change in risk, for example a reduction of fatality rate by 2 in 100,000. If this reduction is valued at 500 NOK, the value of a statistical life is estimated as follows:

$$\text{Value of a statistical life (VSL)} = \frac{500}{\left(\frac{2}{100000}\right)} = 25{,}000{,}000 \text{ NOK}$$

The value of a statistical life is the value of a risk reduction which statistically corresponds to reduction of the number of fatalities by one.

The methodology of scientific research programmes

The first estimates of the value of preventing a road accident fatality were made in the 1950s in Great Britain, Sweden and the United States. These estimates were based on the so called human capital approach. The basic idea of the human capital approach was to estimate the capital value of a human being. This value was usually estimated as the present value of future earnings.

The method had some glaring deficiencies. Children and the retired had no value, since they did not earn anything. Housewives doing unpaid household work also had a value of zero, or, in some studies, a negative value since their consumption had to be supported by others. Besides, the method had no theoretical foundation.

Valuation of life and health as a scientific research programme started by defining the theoretical basis of valuation research. This was done around 1970 by economists Schelling, Mishan and Jones-Lee. They all argued that the only theoretically meaningful approach to the valuation of non-market goods, like life and limb, is the willingness-to-pay approach. The value of any good is indicated by the amount an individual is willing to pay for the good. The more valuable we think something is, the more we are willing to pay for it. Valuation in terms of willingness-to-pay is individual and subjective. No objectively "correct" valuation exists.

But how can we find out whether or not we can trust the results of valuation studies if there is no correct answer to the question asked in these studies? Do we simply have to accept any amount people state? No, as explained later, there are many ways of assessing whether the results of valuation studies can be taken seriously or have to be rejected.

Schelling, Mishan and Jones-Lee placed the valuation of life and limb squarely within consumer theory. This is well-developed branch of economic theory with a long research history. Consumer theory basically seeks to explain how people allocate their spending between different commodities and services. It does so by relying on the assumption that people choose the pattern of consumption that gives them the greatest overall satisfaction. Another way of saying this is that consumers are assumed to maximise utility (the term utility denotes the satisfaction of preferences; to put it colloquially: to maximise utility is to do what you like best).

The assumption of utility maximisation is the core of consumer theory. This theory has the same characteristics as a scientific research programme as defined by Imre Lakatos. Lakatos developed the methodology of scientific research programmes principally as a descriptive theory of how research actually takes place, intended as a conceptual framework for what he called the "rational reconstruction" of the history of science. By rational reconstruction he meant a historical reconstruction of science as the product of rational choices made by researchers working in a field of knowledge. He proposed that the rationality of the choices made by researchers should be judged according to standards defined by the researchers themselves, in particular the standards he labelled "positive heuristic" and "negative heuristic" (see explanation of these terms below).

According to Lakatos, a scientific research programme consists of a hard core, a protective belt, a positive heuristic and a negative heuristic. A programme may, at any point in time, be in a progressive phase or a degenerative phase. The hard core of a scientific research programme consists of basic assumptions made by all researchers working within the programme. The basic assumptions are taken for granted; it is forbidden to raise doubts about them. The basic assumptions are not tested empirically. The hard core is surrounded by a protective belt. The protective belt consists of hypotheses that are derived from the hard core. Hypotheses in the protective belt are tested empirically and can be rejected. Rejection of a hypothesis in the protective belt is in general regarded as undesirable, since any hypothesis in the protective belt is formulated by means of deductive reasoning based on hard core assumptions. Logically speaking, therefore, rejection of a hypothesis in the protective belt casts doubt on the validity of the assumptions forming the hard core. Any finding that, taken at face value, casts doubt on the hard core is called an anomaly. Anomalies, Lakatos argues, do not normally lead to the rejection of the hard core of a research programme. On the contrary, the hard core is normally left intact and research continues as if the anomalies did not exist.

There are usually many interpretations of an anomaly. The positive heuristic of a research programme calls for researchers to increase the empirical content of the programme and, in particular, to develop hypotheses or research methods that will eliminate anomalies. The empirical content of a research programme consists of all observations implied by the hypotheses forming the protective belt, both observations that have been confirmed and observations not yet made. A programme is in a progressive phase when its empirical content increases. A programme enters a degenerative phase when its empirical content no longer increases and when anomalies come to be the normal finding of empirical research. One sign of a degenerative phase is that the anomalies are explained by means of ad hoc hypotheses only, i.e. hypotheses that explain a single anomalous finding, but have no implications predicting novel findings. The negative heuristic calls on researchers not to question the hard core and not to develop ad hoc hypotheses to explain anomalies.

What happens when anomalies become very many? At some point a research programme may be abandoned, but according to Lakatos this does not take the form of a scientific revolution as suggested by Thomas Kuhn. He argues that findings contradicting a theory are, by themselves, not enough to reject the theory. A theory is only rejected when a new and better theory has been developed; better, in the sense that it explains all verified content of a research programme as well as (at least most of) the anomalous findings of that programme. In other words: anomalies cease to be anomalies if a theory is developed that re-interprets them as normal findings.

Application of the methodology of scientific research programmes to valuation research

The methodology of scientific research programmes is highly applicable when trying to reconstruct the history of valuation research.

The hard core of this research is the assumption made in consumer theory that consumers are rational utility maximisers. This assumptions is purely formal; it has no empirical content and merely states that consumer choices can always be modelled as a utility function which is maximised. All one needs to assume to apply this basic postulate, is that preferences can be represented by means of a utility function which has the mathematical properties necessary for a maximum to exist. These properties are very weak and innocent-sounding, boiling down essentially to the requirements that preferences should be transitive and complete. Researchers are at great liberty to make further assumptions about individual utility functions. Thus, as an example, Jones-Lee (1974) made the following assumptions:

1. The individual maximises expected utility (which is a probability weighted utility of a lottery with life and death as potential outcomes).
2. The individual prefers more wealth to less and is financially risk averse (prefers an income received with certainty to an uncertain income).
3. The individual does not want descendants to be exposed to a greater financial risk than himself or herself.
4. At a given level of wealth, the individual prefers to be alive rather than dead.
5. The marginal utility of wealth is greater when the individual is alive than when the individual is dead.

Based on these assumptions, Jones-Lee could deduce that a positive willingness-to-pay for reduced risk of death will exist. He further deduced that willingness-to-pay will be positively related to income and positively related to the level of risk. This example shows how one can use theoretical predictions to assess whether empirical results make sense or not. If you find that willingness-to-pay varies systematically as predicted by theory, results make sense. If you do not find the predicted pattern, interpretation becomes more complicated. It could be that your theory is wrong, but it could also be that the methods were not good enough to uncover the expected pattern. At any rate, this example shows the essential function of theory in willingness-to-pay research: It is to predict a systematic pattern of variation in willingness-to-pay that may serve as reference in assessing whether the results of empirical studies make sense or not.

An important positive heuristic in the early days of valuation research was to develop theory to help identify meaningful patterns of variation in willingness-to-pay. Important theoretical contribution were made around 1980. However, empirical research did not always give results that supported the hypotheses. This was widely interpreted as a problem of research method. Another important positive heuristic in the early days of valuation research was therefore to continuously develop and improve methods for valuation studies.

The progressive phase (1970-1995)

Following the definition of the theoretical foundation of valuation research around 1970, some years passed before the first empirical studies got started. From about 1980, empirical studies were made both in Europe and the United States. Different research traditions developed on the two continents. In Europe, the contingent valuation method was used in all studies until the late 1990s. This method elicits willingness-to-pay by asking respondents to state how much they are willing to pay for a certain change in risk. Several versions of the method have been developed. In the United States, the dominant method was studies of compensating wage differentials. These studies estimated the additional wages paid to compensate for work-place risks.

The first major valuation study in Europe was performed by Jones-Lee and others in Great Britain. The study was originally reported in 1983, but has been published a number of times, including as a chapter in a book by Jones-Lee published in 1989. When the study was in progress, Jones-Lee convened an international research conference in Geneva in 1981. The conference was attended by the leading researchers at the time and several major theoretical contributions were presented during the conference. These theoretical contributions increased the empirical content of valuation research (i.e. they predicted findings of empirical studies).

The empirical content of valuation research increased further as a result of a number of empirical studies. Replications, or near-replications, of the British study were made in Austria, Sweden, New Zealand, Denmark, Switzerland and France. All these studies were reported before 1995. Their findings were, mostly, not very different from the original British study.

The beginning of the end of the progressive phase was an increasing number of anomalous results. These results raised doubts about the validity of contingent valuation estimates of the value of a statistical life. In the United States, there was always a greater scepticism to contingent valuation than in Europe, and in 1993 a critical assessment of the method was published, providing methodological guidelines on how to conduct good contingent valuation studies. A common problem was insensitivity to scope. Insensitivity to scope means that willingness-to-pay does not increase in proportion, or near proportion, to the size of the risk reduction. Thus, it was typically found that people were not willing to pay twice as much for a risk reduction of 4 in 100,000 as for a risk reduction of 2 in 100,000. It was believed that one source of the problem was that people had difficulty in understanding small changes in low levels of risk. The numbers 4 in 100,000 and 2 in 100,000 do not seem to be very different – both are very low numbers.

A French study tried to get around this problem by asking for willingness-to-pay for reductions in the number of traffic fatalities in France ranging from 50 to 5000. The idea was that people would more easily notice the difference between these numbers than the differences between low levels of risk. However, the French were not willing to pay 100 times more for reducing the number of traffic fatalities by 5000 than for reducing them by 50. In fact, they were, on the average, only willing to pay slightly more than 4 times more for reducing fatalities by 5000 than for reducing them by 50. Thus, the problem remained unsolved.

Moreover, it was found that different methods for eliciting willingness-to-pay in contingent valuation studies produced different results, although the methods were equivalent according to economic theory. Things that ought not to make a difference according to economic theory did in fact make a difference. Studies of compensating wage differentials gradually improved, as both new sources of data became available and statistical modelling became more advanced. Yet this progress was not associated with more consistent findings. On the contrary, the range of estimates became bigger, despite progress with respect to data quality and statistical analysis.

By the end of the 1990s, the progressive phase was over. Valuation research entered a new phase.

The struggle between progressive and degenerative tendencies (1995-2005)

The results of contingent valuation studies were, by the end of the 1990s, full of anomalies. Some researchers went so far as to say that all results of such studies were anomalies. Jones-Lee, who had championed contingent valuation in the early phase of valuation research, rejected the method in 1998 and proposed a new method bypassing the need to ask people about changes in low levels of risk.

From about 2000, a number of valuation studies started to use the stated choice design. Respondents were asked to choose between two alternatives, in most cases two roads. The roads differed with respect to, for example, travel time, number of accidents and toll charges. Respondents were asked to choose which road to take; the attributes of the two roads were then modified and respondents asked to choose once more. Each respondent would typically make 5-10 choices. The valuation of the non-monetary attributes was obtained by analysing the choices made, usually relying on random-utility functions (i.e. functions respondents were assumed to maximise, but that would have a residual terms since analysts did not know all factors influencing utility).

In the United States, Peter Dorman launched a strong criticism of studies of compensating wage differentials in 1996. He argued that all these studies were flawed and should be rejected. He concluded that one should abandon monetary valuation of life and limb entirely. The leading proponent of wage-risk studies in the United States, Kip Viscusi, dismissed Dorman's criticism. Many studies of compensating wage differentials have been published since 1996. Dorman was not successful in his attempt to bring this research to an end. He may well have been right in much of his criticism, but, as Lakatos pointed out; criticism per se is rarely enough to overturn a scientific research programme. You have to offer something better. Dorman did not offer any alternative. He simply said: "Stop doing this", without saying what one should rather do.

Stated choices were initially thought to be superior to contingent valuation, for example because people were not asked about changes in low levels of risk. However, anomalies soon turned up in stated choice experiments. Lexicographic choices were common. An individual chooses lexicographically if he or she always prefers the alternative that is best with respect to one of the attributes, and ignores the others (for example, always chooses the safest road). Many respondents made inconsistent choices. This means that at stage N in a sequence of choices, they preferred an alternative implying a valuation which was inconsistent with a choice made at stage N – 1 of the sequence. Only about 10-20 percent of respondents made choices that were fully consistent with economic theory.

Research was a struggle between progressive and degenerative tendencies. The launching of new methods represented the progressive element; the repeated finding of anomalies represented the degenerative element. The anomalies did not go away, they merely took new forms.

Meanwhile, the theoretical foundation of valuation research was undergoing a rapid transformation.

The protective belt becomes almost all-inclusive (2000-2010)

The early theoretical contributions to valuation research made clear predictions about empirical results. As an example, Jones-Lee predicted that willingness-to-pay would increase as the level of risk increased. Finding the opposite would falsify his theory. It did not take long, however, before more complex models were developed and predictions became ambiguous. The extremely complex model proposed by Dehez and Drèze is an example. Here are the predictions of this model:

1. If an individual does not have life insurance or an annuity, and if the marginal utility of money is greater when alive than when dead, willingness to pay will increase when risk level increases.
2. If the individual has optimal life insurance and annuity at actuarially fair rates, willingness to pay is independent of the level of risk.
3. If the individual holds life insurance and annuity at less than actuarially fair rates, willingness to pay will increase as the level of risk goes down.
4. If the individual holds life insurance and annuity at more than actuarially fair rates, willingness to pay will increase as risk level increases.
5. If the individual has life insurance and annuity and the terms of the contracts are adjusted as risk level changes, willingness to pay will increase as risk level decreases.

One could say that they hedge their bets. Everything is possible; that willingness to pay does not depend on risk level, that it increases with risk level, or that it decreases with risk level. None of these findings is ruled out theoretically. No matter what you find, it has theoretical support – unless, that is, that you can collect detailed data on the insurance coverage of respondents. But even if such data are available, it may be difficult to determine if insurance is actuarially fair or not. One would normally expect insurance to be less than actuarially fair, but in some countries tax rebates for life insurance may make insurance contracts close to actuarially fair.

In short: It is in practice almost impossible to falsify the hypotheses proposed by Dehez and Drèze. Their contribution was the start of a series of theoretical contributions that have reached the point where almost no finding contradicts theory. Perhaps the most consequential contribution is the directionally bounded utility function introduced by Amiran and Hagen (2003, 2010). This utility function predicts insensitivity to scope, turning what was long regarded as a major anomaly in valuation research into a theoretically expected finding, perfectly consistent with rational utility maximisation. This has revolutionary implications. If insensitivity to scope is to be expected, all the efforts that have made to develop methods to increase sensitivity to scope look like a complete waste of time. Almost any finding must be taken seriously if one takes directionally bounded utility functions seriously.

One could, to be sure, try to ascertain whether respondents do indeed have a directionally bounded utility function or a utility function which is not directionally bounded. However, it is by no means clear how to do this, and a wide range of choices can be consistent with a wide range of mathematical forms of a utility function. It seems quite likely that the data that could realistically be obtained would be inconclusive.

Today, theory has come close to an immunising stratagem. An immunising stratagem, a concept introduced by Karl Popper, is a reformulation of a scientific theory so as to make it immune to falsification. Popper would say that such a reformulation makes the theory unscientific, since he regarded only theories that could be falsified as truly scientific. What seems clear, is that bringing the theory close to an immunising stratagem undermines its function in research. If no result can be ruled out on theoretical grounds, it is no longer possible to refer to theory to support or reject an empirical funding. Theory no longer discriminates between meaningful and meaningless findings.

A hard core in dissolution? (2005-2015)

There is little doubt that the original ambition of valuation research was to find a single value of a statistical life that could be applied uniformly. The many references to allocative efficiency made in early contributions attest to the importance given to this objective. Allocative efficiency can only be attained if the same value of a statistical life is used in all sectors of society. If one allows much more to be spent on saving life in one sector than in another, one may in principle save a larger number of lives by transferring spending from the "expensive" sector to the "cheap" sector.

As research has produces an ever widening gap in estimates of the value of a statistical life, the objective of finding a single value that can be applied universally has been toned down. Many researchers have argued that a single value does not exist. We know, for example, that age and income are likely to influence willingness to pay. In recent years, some researchers have started to ask whether one should allow the value of a statistical life to vary, rather than using a single, uniform value. The issue then becomes how much variation to allow for and what sources of variation would be legitimate. Both Jones-Lee and Viscusi have argued that income is legitimate: it is entirely appropriate to treat rich people's lives as more valuable than poor people's lives. They both argue that doing so is more consistent with the theoretical foundation of cost-benefit analysis than using a single, uniform value of a statistical life.

At this point a full circle has been travelled. Research set out to find a single, universally applicable value of a statistical life. It was quickly realised that such a value does not exist. Hypotheses were developed to predict systematic variation in valuation. The underlying idea was not necessarily that the value of a statistical life should also vary; it was rather that if one found variation making sense according to economic theory, more trust could be placed in findings than if one did not find such a variation.

The expected variation was only partly found. Besides, a lot of variation attributable to sources that ought to be irrelevant according to economic theory was found. As this continued in study after study, researchers turned their attention to theory once more and embarked on reformulating it to enable the anomalies to be interpreted as normal findings. This has been so successful that probably quite few findings would now be regarded as anomalous.

The huge variation in estimates of the value of a statistical life was thus transformed from a problem into something which is to be expected. Doubts have been raised about the prescriptive ideal of applying a uniform value of a statistical life in cost-benefit analyses.

Today, therefore, little remains of the original research programme that inspired valuation research. That programme has travelled full circle and come to an end in the sense that the hope of finding a single value of a statistical life has been given up and even the ideal of allocative efficiency, as traditionally understood, is being questioned.

We will nevertheless ask: Where do we go from here? What are the prospects for valuation research?

Can meta-analysis create order in chaos? (2000-2015)

A number of meta-analyses of value of life studies have been reported during the last fifteen years. The report presents these analyses and discusses whether they have been able to make sense of the huge diversity of estimates of the value of a statistical life.

Most of the meta-analyses are somewhat simplistic and do not fulfil the methodological standards for high-quality meta-analyses. The analyses do, for example, not apply optimal statistical weights, nor do most of them test for the possible presence of publication bias. Some of the meta-analyses have tried to sort primary studies according to study quality.

Norwegian researchers Ståle Navrud, Henrik Lindhjem and Nils Axel Bråthen, co-operating with the French economist Vincent Biausque, have made the most comprehensive meta-analysis reported so far. It has been published in several rounds; the most recent included 931 estimates of the value of a statistical life. Meta-regression analyses were performed in order to explain variation in estimates; these analyses were made both for the entire sample and for a subsample of the studies that were classified as methodologically best.

The factors that were found to have the largest influence on the value of a statistical life were income (higher income, higher value), size of the risk change (larger change, smaller value) and whether safety was provided by means of a public good or a private good (lower value for public goods). It may strike readers as surprising that the value of a statistical life was lower when the change in risk was large than when it was small. This is attributable to insensitivity to scope, as shown by the numerical example given below:

Risk reduction	Willingness-to-pay	Value of a statistical life
$1 \cdot 10^{-5}$	400	40,000,000
$2 \cdot 10^{-5}$	500	25,000,000
$5 \cdot 10^{-5}$	600	12,000,000
$10 \cdot 10^{-5}$	800	8,000,000

Considerable publication bias has been found in the value of life literature. Studies that have adjusted for publication bias indicate that by doing so, the value of a statistical life drops to about one third of the un-adjusted value (e.g. from 9 million to 3 million).

Although some of the meta-analyses explain most of the variation in the value of a statistical life, it is apparent that residual terms are very large, in particular at the ends of the distribution. Moreover, most meta-analyses do not recommend a best estimate of the value of a statistical life. The analyses also remain silent on the topic of whether a variable value of a statistical life should be applied.

Can more promising methods be found?

Valuation research has so far not been able to produce very precise estimates of the value of a statistical life. Even within the same study, estimates often vary be a factor of 10 or more. Can we think of other methods that might produce more precise estimates?

The report reviews a few options. One of them is to use Quality Adjusted Life Years (QALYs) as a starting point. A QALY is a numerical scale for quality of life related to health state. By convention, death is given the value of 0 and perfect health the value of 1. States of reduced health as assigned values between 0 and 1, closer to 0 the worse they are. If a health state involving a loss of, for example 0.02 QALYs is valued at 50,000, the idea is that one may scale this up to a value of a statistical life year of $1/0.02 \cdot 50{,}000 = 2{,}500{,}000$.

There are two big problems with this approach. In the first place, there is no universally accepted method for obtaining QALYs. On the contrary, quite many scales can be found in the literature and they do not agree. A specific health state will therefore not have the same QALY score according to all scales. This means that estimates of the value of a statistical life may differ substantially, depending on which QALY scale is used to scale up the value applying a slight reduction of health to the value of a statistical life. In the second place, QALYs rely on utility functions with very restrictive properties; indeed much more restrictive than the utility functions normally assumed in conventional value of life studies. The QALY approach is therefore not very promising.

Another approach discussed is the capability approach introduced by Amartya Sen. The core of this approach is that to obtain an acceptable quality of life and standard of living, an individual must possess certain capabilities, such as access to clean water, literacy, housing, and so on. Sen has proposed this approach as an alternative to subjective well-being, noting that even people who live in great hardship often report a high level of subjective well-being. The capability approach resembles the "social indicators movement", which proposed to measure how well a society is taking care of its citizens by using social indicators like literacy rates, access to water and toilet facilities, and so on. Indicators like these are probably best suited to low-income countries. They are completely inconsistent with the theoretical foundation of valuation research by being fully paternalistic. The capability approach is hence judged as irrelevant to valuation research.

The third approach discussed is to develop empirical utility functions based on surveys of subjective well-being (happiness surveys). Such surveys have been made in many countries. By studying the relationship between reported subjective well-being and, for example, income, functions can often be fitted that have the same characteristics as utility functions as usually defined in economic theory. Interpreting such empirical functions as empirical utility functions is, however, still controversial among economists, although the idea seems to be gaining increasing support. Another problem is that different data sets give rise to different functions, that may not imply the same valuation of a statistical life.

None of these three approaches would therefore seem to be clearly superior to the conventional designs used in valuation research so far.