Estimating the crime value of a QALY: A methodological review.

Emily Bulman[3], Judith Covey[2], Graham Loomes[1], Angela Robinson[1], Michael Spackman[3]

1 University of East Anglia

2 Durham University

3 NERA Economic Consulting

Corresponding author;

Angela Robinson,
Senior Lecturer in Health Economics,
School of Medicine, Health Policy & Practice,
University of East Anglia
Norwich
NR4 7TJ

01603 593620

angela.robinson@uea.ac.uk

INTRODUCTION

Quality adjusted life years (QALYs) are commonly used in health economics and their use required in submissions to NICE. Other government agencies with responsibility for health and safety expenditure use a more traditional cost-benefit approach to valuation but the Home Office has recently adopted a 'QALY-based' approach to valuation. The prospect of 'exporting' QALYs to other areas clearly raises a number of methodological issues which we set out to explore in this paper.

The Home Office publishes figures for the economic and social costs of crime against individuals and households. The first issue was published in 2000 and the current issue in 2005 (hereafter described as HOOR 30/05).[1] These costs are estimated as the sum of costs "in anticipation of crime", costs "as a consequence of crime", and Criminal Justice costs "in response to crime". Costs "as a consequence" of crime include elements such as property loss or damage, lost output and health service costs, but the largest component in aggregate is "Physical and Emotional Impact on Direct Victims".

The purpose of the present study is to reconsider those values in the light of more recent research/developments and to identify areas for future research that would help inform the possible adjustment of those values. HOOR 30/05 explains that the new method adopted in 2005 was to translate the physical and emotional impacts of each category of violent crime into losses of QALYs. A monetary value for the QALY was derived by translating into QALYs the physical impacts of a transport injury for which a WTP monetary value had been obtained. The effect of this change of methodology was to produce large increases in the values assigned to the physical and emotional impacts of sexual offences and common assault and a large decrease in the value associated with wounding.

What will be of particular interest to health economists is that the value of a QALY underlying these estimates is considerably higher than that currently used in health. On the other hand, the Home Office have some concerns that certain 'intangible' and 'contextual' features related to

---

[1] The current issue is "The economics and social costs of crime against individuals and households 2003/04", Home Office Online Report 30/05, http://www.homeoffice.gov.uk/rds/pdfs05/rdsolr3005.pdf .

crime are not currently taken into consideration and wonder if the value ought to be inflated according to some 'crime premium'. This paper sets out to review the existing QALY-based methodology and those possibly important factors that the current methodology does not cover, such as the nature or perceptions of the risks involved and other 'contextual' factors.

REVIEW OF CURRENT PROCEDURE

The current approach is based on work carried out by Dolan et al (2003).[2] The Dolan et al team (which included an author of this current paper- GL) were commissioned to review the techniques for estimating the intangible victim costs of crime, they were asked to focus on the physical and emotional impacts on victims themselves[3]. To that end, they considered three possible ways of arriving at monetary values for (the prevention of) such impacts for each category of violent crime:

- Method 1: Direct elicitation of money values, either from the 'revealed preferences' (RP) or from the 'stated preferences' (SP) of (a representative cross-section of) the population.
- Method 2: Transfer of (possibly modified) money values from other UK public sector policy areas.
- Method 3: Estimation of quality/disability adjusted life year (QALY/DALY) indices and conversion via some monetary value for a QALY.

The current Home Office cost of crime estimates are based on the third of these approaches. That is, for each type of offence, the DALY/QALY loss entailed by each category of possible physical or psychological injury is weighted by the prevalence of that injury in order to produce an estimate of the expected QALY loss associated with each crime type. DALY scores were computed for each category of offence involving violence to the person. Data from the British Crime Survey (BCS) was used to map categories of crime to prevalence of different levels and durations of physical injury. Data from a review of research on victims of crime was used to map crimes to prevalence of different categories of psychological disorder; and to the duration of

---

[2] Throughout this paper references to "Dolan et al" refer to this 2003 Report.

[3] Given this remit, our consideration of other factors not currently costed is not intended as a criticism of their study.

physical and psychological impacts. These health states of victims were then mapped to QALYs, using weights taken from a variety of research sources, including use of the DALY, and a Dutch study for psychological impacts.

This expected QALY loss is converted into a money value by applying a value of a QALY of approximately £81,000 in 1997 prices (that is, just under £109,000 in the 2003 prices cited in Table 2.1 of Section 1 of HOOR 30/05). This money value of a QALY was derived by taking a particular road accident injury (known as Injury W), calculating the QALY loss associated with that injury (estimated at 0.037 QALYs) and linking that to a stated preference value for preventing that injury (rounded to £3,000 in 1997 prices, taken from Carthy et al. (1999)). It is important to note that the estimate taken from Carthy et al was a weighted average of the WTP *and* WTA values generated in that study. Using WTP values alone would have resulted in a significantly lower estimate for the value of a QALY[4]. It is also interesting to note that the approach to valuing a QALY that appears in the appendix of the Mason et al paper (2008) in Health Economics is also based on mapping road traffic injuries onto QALYs and WTP valuations from Carthy et al, though they came up with rather different estimates.

Whilst we conclude that the chosen method has much to recommend it, there are some doubts about the robustness of the particular estimates derived from it and are currently being used. We are keen to stress here that, in outlining these caveats (a number of which were raised in their original report), we are in no way questioning the quality of the original research undertaken.

It could be argued that one advantage of the method used is that estimates for all of the crimes involving non-fatal physical and emotional damages are based on injury descriptions that are reasonably representative of what victims actually suffer (rather than rely on the hypothetical choices of perhaps ill informed respondents to surveys). Such an approach, by its very nature, also involves giving attention to both levels *and* durations, whereas respondents in SP surveys appear liable to be insufficiently sensitive to these dimensions. There are, however, always disadvantages and limitations of any approach taken and we outline several below.

[4] This is not a criticism of the approach Dolan et al adopted- the Carthy et al paper sets out the argumnents in favour of using WTP/WTA 'hybrid' values.

First, the various injuries typical of those suffered by victims of crime had to be given QALY scores using existing systems that had been developed for rather different purposes and were far from ideally suited to this task. Moreover, mapping between the various descriptor systems in order to arrive at QALY estimates inevitably requires a degree of subjectivity and no formal tests of inter-rater reliability were possible within the timescale of that research.

Second, having arrived at a particular set of QALY weights, the next step involved finding a way of converting them into money values. In order to convert QALY measures into monetary values, some 'rate of exchange' needs to be determined. However, the derivation of any such monetary value of a QALY – henceforth, MVQ – is neither theoretically nor empirically straightforward. The method used was to assign a QALY score to road injury W and link that to a rounded weighted average value drawn from a particular survey. Thus, much hinges on a single QALY score linked to a single value from just one study.

Third, the measures take no explicit account of risk. The EQ-5D 'tariffs' used to index health states are based on TTO responses with no consideration of risk, while the Global Burden of Disease (GBD) weights are derived from expert judgments made without reference to risk attitudes. However the policy reality is that many measures are intended to reduce the (usually very small) risk people face of suffering these injuries, rather than aiming to prevent victimisation that would otherwise occur with certainty.

Fourth, there may be intangible costs that are not likely to be amenable to such measures. If, for example, the consequence of someone being assaulted or raped is that other members of the household experience some degree of distress and/or raised levels of anxiety, this will not be captured in a QALY-type framework (whereas it could- at least in theory- be captured in direct WTP valuations). There could also be 'contextual' effects that make crime-related injuries different from road traffic injuries even if the health effects are apparently identical. This raises the question of whether or not the estimated value of a QALY based on road transport injury W ought to be multiplied up by some 'crime premium'. Whilst the second half of the paper addresses the issues of risk perceptions and possible contextual effects in considerable detail-we deal with the first two points in turn here.

MAPPING BETWEEN INJURIES AND HEALTH STATES

As above, an approach adopted by Dolan et al (2003) and Mason et al (2008)[5] was to 'map' from those injuries for which WTP values already exist onto other states for which utility (or disutility) estimates may be calculated and, hence, allow MVQ to be estimated.

The potential problem here is that there may be no unique mapping between those states for which health state utility indices exist and those for which WTP values exist. Moreover, starting with a different injury for which WTP values exist may lead to different MVQ estimate. We illustrate this point by detailing a small 'mapping' exercise which set out to test whether 'experts' mapped the non fatal injuries valued directly in Carthy et al (1999) (injuries W and X) onto different EQ-5D and SF-6D states.

Non-fatal road traffic injuries W and X were first broken down into their 'in hospital' and 'after hospital' components (both are temporary injuries- X being more severe than W, lasting 18 months, rather than a few weeks). Respondents were asked to map the each component onto the EQ-5D and SF-6D descriptive systems. The short questionnaire was administered to members of the Health Economics Group (HEG) at UEA at the end of one of their regular research meetings. Nine members of the group returned completed questionnaires. Only the data from the EQ-5D are reported here.

The nine raters mapped the 'in hospital' component of injury W onto nine different EQ-5D states with a range of 'tariff' utility values of between 0.364 (for condition 21321) to 0.796 (for condition 11121). W 'after hospital' was mapped onto 6 different EQ-5D states with a range of utility values of 0.516 (for 22222) to 0.76 (for 11221). The nine raters mapped the 'in hospital' component of injury X onto 9 different EQ-QD states with a range of utility scores of 0.002 (for 32222) to 0.727 (for 21121). The 'after hospital' component of injury X was mapped onto 6 different EQ-5D states with a range of utility scores of 0.516 (for 22222) to 0.760 (for 11221). After multiplying the utility losses by the duration of each component of the injury, QALY

[5] This refers to work that appears in the appendix of that paper- rather than the main body of the modelling work.

losses were calculated. The QALY loss associated with injury W ranged from 0.083 to 0.165 and for injury X ranged between 0.370 and 0.648.

Table 1 shows the range of tariff utility scores our mapping exercise yielded;.

	Min	Max
W in hospital	0.364	0.796
W after hospital	0.516	0.76
X in hospital	0.002	0.727
X after hospital	0.516	0.76

These values then allow the QALY loss associated with an injury to be calculated. Combining this with the WTP value for this injury allows the WTP per QALY to be estimated (by dividing the WTP value by the QALY loss).

Table 2 reports the WTP and WTA estimates for injuries W and X reported in Carthy et al (1999);

	Mean	Median
WTP for W	£1,733	£1,000
WTA for W	£11,952	£3,000
WTP for X	£5,258	£3,000
WTA for X	£33,746	£10,000

Applying the mean WTP for W values to the mean QALY losses derived in our mapping exercise yields a mean estimate of £15,839 per QALY gained. Applying the *median* WTP for W value yields an estimate of £9,140 per QALY gained. Using a 'hybrid' value of WTP/WTA estimates[6] - see Carthy et al (1999)- results in an estimate of £46,926 per QALY using injury W.

[6] Carthy et al discuss a range of weighting systems for combining WTP and WTA values- this illustrative example uses an 2/3 WTP + 1/3 WTA weighting. The implied WTP per QALY gained will clearly depend on the weighting system adopted.

Turning to injury X, applying the mean WTP values to our estimated QALY loss data results in mean estimates of £11,154 per QALY gained. Using *median* values results in estimates of £6,364 per QALY gained. Again, using a hybrid of WTP/WTA yields mean estimates of £31,267 per QALY gained.

Whilst no weight may be attached to these numerical estimates and we have no reason to suppose that the judgments of our raters are superior to anyone else's, the exercise highlights two important points in assessing the validity of mapping from injuries for which WTP values exist onto those for which we have QALY (or DALY) measures. First, we have demonstrated that the injury descriptions may be mapped differently by different raters. Second, we have demonstrated that using the same raters but starting with different injuries for which WTP values exist can result in substantially different estimates: the estimates derived from the more minor injury W were almost 50 per cent higher than the corresponding estimates derived from the more severe injury X.

An additional issue, though not related to mapping per se, is that using different measures of central tendency of WTP values from the same study- or even different conventions for weighting WTP and WTA values- can result in significantly different estimates of MVQ. Again, this is not a criticism of the work that was carried, but raises the point that if we are to 'export' QALYs to other sectors (and resources were to be allocated on that basis) it will be important to ensure a consistent methodology is used in deriving MVQ.

THE 'RATE OF EXCHANGE' BETWEEN QALYS AND MONEY VALUES

We turn now to the more general- and conceptual- issue of establishing a rate of exchange between QALYs and money values. There are, of course, well known criticisms of the QALY model itself which we do not need to rehearse for this audience. But, even in cases where the basic QALY assumptions hold to an acceptable degree, theory provides no guarantee that any individual will have a single MVQ.

Two recent substantial studies have investigated the feasibility of generating a reasonably stable MVQ. Pinto Prades et al (2009) found considerable variability, depending on the severity and duration of the health states used and also depending on the order in which questions were asked:

their mean estimates of the MVQ varied from less than €5,000 to more than €110,000. A feasibility study (Chapter 6 in Donaldson et al., 2008) also showed how mean MVQ might range from less than £20,000 to more than £250,000 depending on the scenarios used.

In addition, a number of health economic studies have previously elicited utility values alongside WTP responses and found no linear relationship between utilities and WTP values, and, hence, no unique MVQ. Bala et al (1998) found 'no significant correlation' between WTP and SG values. Lundberg (1999) found WTP values to be correlated with a disease specific measure of health the (dermatology life quality index), but not with utility scores or SF-36 scores. Franic, Pathak and Gafni (2005) found that responses from VAS, SG and TTO were poorly correlated with WTP, leading them to question whether a stable WTP per QALY could ever exist. Gyrd-Hansen (2006) found little correlation between WTP and TTO values. Other studies have reported 'some evidence of convergent validity' between SG and WTP (Obrien and Viramontes, 1994). The emerging picture, however, suggests that WTP responses are not well correlated with utility scores.

The results reported above appear to raise serious doubts as to whether a stable and robust MVQ *could* ever exist- which, of course has just as many implications for establishing a health MVQ as one in the field of crime.

CONTEXTUAL FACTORS

We turn now to the question of whether there are any grounds for arguing that crime is 'special' and, hence, that MVQ used by the Home Office ought to be multiplied up by some 'crime premium'. Given the remit of our work for the Home Office, the purpose of our review of contextual factors was to determine whether crime risks might be valued differently from road risks (recall that the value of £81k was based on WTP/WTA values for road traffic injury W). But the research is clearly also relevant to the question that more likely interests health economists: Ought a crime QALY be valued significantly more highly than a health QALY?

Contextual factors may cause the potential victim's or the public's valuation of risk reductions in one hazard context to differ from their valuation in another context. Influential research was conducted by Paul Slovic and his colleagues in the 1980s to identify the characteristics of

hazards that may trigger especially high or low levels of alarm, fear or outrage in potential victims and the general public (e.g. Slovic et al. 1980; Slovic, 1987). The findings on hazard characteristics are well summarised and developed in a Department of Health publication of 1997 on "Communicating about risks to public health". Illustrative "fright factors" are shown in Box 1.

Box 1: Fright Factors Listed by the Department of Health (1997, p5)

Risks are generally more worrying (and less acceptable) if perceived:

1. to be **involuntary** (e.g., exposure to pollution) rather than voluntary (e.g., dangerous sports or smoking)
2. as **inequitably distributed** (some benefit while others suffer the consequences)
3. as **inescapable** by taking personal precautions
4. to arise from an **unfamiliar or novel** source
5. to result from **man-made, rather than natural** sources
6. to cause **hidden and irreversible** damage, e.g., through onset of illness many years after exposure
7. to pose some particular danger to **small children or pregnant women** or more generally to **future generations**
8. to threaten a form of death (or illness/ injury) arousing **particular dread**
9. to be **poorly understood by science**
10. is subject to **contradictory statements** from responsible sources (or, even worse, from the same source)

Source Department of Health (1997)

A key word/citation search was conducted by one of the authors (JC) on Web- of Knowledge and research papers included in the review if they satisfied the following inclusion criteria;

- Research which has compared people's preferences for risk reductions/ life-saving interventions between different types of hazard contexts

And/Or

- Research which has examined the relationship between risk perception factors and people's preferences for risk reductions/ life-saving interventions

Papers were excluded if the hazards studied could not differentiate between crime risks and road risks (latency of hazard and the affect on future generations).

This literature search yielded a total of twenty relevant papers. The key method used (including sample sizes, P) are summarised in the Appendix and the key findings discussed below.

A range of methods were used in these studies to gain insights into people's attitudes towards different types of risk reduction. The most common type of method was based on eliciting people's preferences for saving specified numbers of lives or reducing risks from different types of hazards[(i,ii,v,vi,vii,ix,x,xi,xii,xiii,xviii,xix)]. For example in the earliest of the studies reviewed, Jones-Lee et al. (1985)[(i)], people indicated their preferences for saving 100 lives from either motor accidents, heart disease, or cancer. Five of the studies[(vii,ix,x,xiii,xviii)] attempted to quantify the strength of people's preferences by using 'matching' questions to estimate the extent to which people are willing to trade-off saving lives from one type of hazard against saving lives from another type of hazard. This approach is similar to the person trade-off method which has been used to elicit people's social preferences for health care interventions.

A smaller number of studies have used monetary valuation methods to gain insights into people's relative valuations for life-saving or risk reducing interventions that would benefit society as a whole[(iii,iv,viii,x)]. This includes studies which have asked people to allocate a fixed budget between different life-saving interventions[(iv,viii)].

Only the final group of five studies have elicited people's purely self-interested preferences for risk reductions that would benefit only their household or them personally [(xiv,xv,xvi,xvii,xx)]. This group includes the study by Carlsson et al. (2004)[(xiv)], which elicited people's willingness to pay higher fares for either safer taxi journeys or safer air travel.

Turning to the specific types of hazards investigated, motor/car accidents have been included in the most studies[(i,ii,iii,iv,vii,viii,x,xiii,xiv,xvii,xix,xx)]. Other types of transport risks to air passengers, cyclists, pedestrians and rail users[(ii,iii,vii,iv,ix,x,xiii,xiv,xviii,xix)], cancer[(i,ii,iv,ix,xi,xv,xvi,xx)], and environmental risks such as air/ water pollution and pesticide risks[(ii,iii,v,x,xi,xvii)], have also been investigated in quite a few studies. Notably, homicide has only been investigated in one study by Baron (1997)[(viii)].

The most consistently highly prioritised hazard is cancer. Reducing the risk of cancer is valued more highly than reducing the risks of other diseases with similar impacts[xv], heart disease[i,ix], motor accidents[i,ii,xvi,xx]. There is also evidence to suggest that in some cases priorities are given to reduce risks where the victims are less responsible for their own safety. For example, the life of a pedestrian is valued more highly than the life of a car driver[xix], and the life of a rail passenger killed in a collision accident is valued more highly than the life of a passenger who has fallen from a station platform under the influence of alcohol[xviii].

The relative priorities found are not however always consistent across studies. For example, whilst some studies show higher priorities for reducing air pollution risks[x] others show higher priorities for reducing traffic/ car accident risks[xvii]. Opposing patterns of preferences have also been found for aviation and automobile risks[iv, xiv]. These inconsistencies may of course be attributable to features of the study design such as the samples recruited and elicitation methods used.

A range of different characteristics which might underpin people's preferences have been explored in twelve of the studies[iii,iv,v,vi,vii,viii,xi,xii,xiii,xvi,xvii,xix]. The characteristics include the perceived voluntariness of being exposed to the hazards, whether people were to blame for their deaths or responsible for their own safety, the degree of control people have over their own safety, how much dread or uneasiness is associated with thinking about the hazard, how much knowledge either people themselves or experts have about the risks, how many deaths each year are caused by the hazard (prevalence), the age of the victims, and the perceived personal risk or personal exposure to the hazard.

The results for each of these characteristics are summarised in Table 3 below. Significant relationships between risk characteristics and priorities are shown as either positive (+) or negative (-). Blank cells that are not shaded grey indicate that the characteristic was measured in the study, but the relationship with priorities was not statistically significant. Grey shaded cells indicate that the characteristic was not measured.

The results show that positive relationships were typically found for dread, prevalence and personal exposure. Hence, hazards that made people feel uneasy, caused more deaths per year

and felt a high risk from tended to be assigned higher priority. On the other hand negative relationships were typically found for perceptions of voluntariness, blame/responsibility, controllability and age. Hence, hazards that were perceived to be highly voluntary, controllable and where victims were older or more to blame for their deaths tended to be assigned lower priority.

There were exceptions however. One of the two studies that investigated the blame issue did not find a significant result (although this might have been explained by a high correlation with controllability)[xi]. Also although voluntariness was measured in five studies the relationship was only significantly negative in two[vii xiii] and it was positive in one[xvi]. Control also emerged as significantly positive in one study. Notably both of these studies [(xvi,xvii)] asked participants for estimates of their private willingness to pay for reductions in mortality risks, which suggests when people are answering on the basis of the own self-interested preferences they prioritise reducing risks that they feel can be controlled more easily.

We found a paucity of literature directly relevant to crime and concluded that extrapolating results from research done on other risks was not ideal. After considering how the various factors relate to the area of crime and the evidence base for the existence of such factors (space does not permit this detailed discussion here), we concluded that there was a lack of robust evidence on which to estimate the extent to which a range of potential intangible costs might add to the current QALY (or DALY) based measures.

HESG members may, however, be more interested in considering how the results presented here might inform a) whether health QALYs ought to be valued lower than crime QALYs and b) whether all health QALYs ought to be valued the same.

Table 3: Significant relationships between risk characteristics and priorities for risk reducing/ life-saving interventions

	Iii	**Iv**	**v**	**Vi**	**vii**	**viii**	**Xi**	**Xii**	**Xiii**	**xvi**	**xvii**	**xix**
Voluntariness/ choice					-				-	+		
Blame/ Responsibility					-							
Controllability					-		-		-		+	
Dread	+	+							+	-	+	
Knowledge		-								+		
Prevalence			+			+		+	+			
Age				-					-			-
Personal Exposure/Risk	+		+				+		+	+	+	

Note: Blank cells that are not shaded grey indicate that the characteristic was measured in the study but the relationship with priorities was not statistically significant. Grey shaded cells indicate that the characteristic was not measured in the study

CONCLUSION

We concluded that the current methodology for valuing the intangible impacts of violent crime is based on high quality research, but conceptual and empirical problems exist that must call into question the robustness of the estimate currently used. We found the main vulnerabilities stem from the mapping of a transport injury into QALYs to derive a monetary value for the QALY and the mapping of crime injuries into QALYs. While the QALY plays a major role in helping to steer health policy, attempts to monetise it face serious difficulties. Though not covered here, there are clearly other intangible impacts of specific crimes, beyond those on the direct victim that would not readily be covered in a QALY-type-framework.

Another challenge is the effects of contextual factors, such as controllability and prevalence, but our examination of these factors, in the context of the psychological literature, suggests that their effects are fairly small.

Further research should be expected to provide a more robust foundation for the basic methodology, whether it is based on the QALY/DALY, or on the QALY/DALY plus other factors, or on direct WTP valuation of a selection of crime injuries. We conclude by briefly outlining what the various strands of this research may cover;

- **Strand 1: Examination of the robustness of QALY/DALY/SF scores** This work would build on the small mapping exercise highlighted above and establish more clearly the potential for achieving consistent mapping between injury descriptions and QALYs or other health state scores. This would need to be extended by taking a set of descriptions based on BCS and other data and asking a group of raters to map these into different systems, and on that basis derive QALY/DALY scores. The spreads and central tendencies of the resulting estimates could be examined and compared with the figures derived in Dolan et al (2003).

- **Strand 2: 'Within-crime' relativity judgments by members of the public** This would provide empirical evidence for the relative valuations given to different types of injuries by establishing direct observations of people's relative valuations of different crime injuries. The method used would ideally be amenable to the inclusion of psychological harm – Acute Stress Disorder and Post Traumatic Stress Disorder – and compare these with physical injuries.

- **Strand 3: Relativities between crime injuries and transport injuries** In at least some cases, it is plausible that very similar sets of physical and/or psychological injuries might result from criminal attacks as result from transport accidents. This allows the possibility, possibly alongside the work in Strand 2, of asking respondents to make direct judgements about the relative weight to be assigned to preventing injuries sustained as a result of crime as compared with the same injuries sustained in transport accidents.

- **Strand 4: Direct WTP measurement of crime-related injuries**. This is arguably the most radical departure from the current approach and no stated preference approach (for example, WTA/compensation, WTP for risk reduction etc) would be unproblematic in this context.

We look forward to hearing the views of HESG members on what they consider to be the most promising avenues for future research.

ACKNOWLEDGEMENTS

We are grateful to Tessa Peasgood for providing additional details of the procedures used in the original study. We are aware that our brief outline above does not do justice to the very detailed and thorough work that was undertaken. We thank Phil Kinghorn for assisting with the small mapping exercise carried out at UEA and to members of HEG for providing mappings. This study was funded by the Home Office.

REFERENCES

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

Baron, J. (1997) Confusion of relative and absolute risk in valuation. *Journal of Risk and Uncertainty*, **14**(3), 301-309

Carlsson, F., Johansson-Stenman, O., & Martinsson, P. (2004) Is transport safety more valuable in the air? *Journal of Risk and Uncertainty*, **28**(2), 147-163

Carthy T., Chilton S., Covey J., Hopkins, L., Jones-Lee M., Loomes G., Pidgeon N. and Spencer A. (1999) On the contingent valuation of safety and the safety of contingent valuation: part 2 – the CV/SG 'chained' approach. *Journal of Risk and Uncertainty* **17,** 187-213

Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N. and Spencer, A. (2002) Public perceptions of risk and preference-based values of safety, *Journal of Risk and Uncertainty,* **25**(3), 211-232

Cookson, R. (2000) Incorporating psycho-social considerations into health valuation: an experimental study. *Journal of Health Economics, 19*(3), 369-401

Covey, J. A. (2001). People's preferences for safety control: Why does baseline risk matter? *Risk Analysis, 21*(2), 331-340

Covey, J., Robinson, A., Jones-Lee, M. and Loomes, G. (2006) T616: Value of preventing a fatality. Rail Safety and Standards Board

Cropper, M. L., Aydede, S. K., and Portney, P. R. (1994) Preferences for life saving programs - how the public discounts time and age. *Journal of Risk and Uncertainty, 8*(3), 243-265

Department of Health (1997) "Communicating about risks to public health: Pointers to good practice", Department of Health
http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4006604

Dolan, P., Loomes, G., Peasgood, T. and Tsuchiya, A. (2003) Estimating the intangible victim costs of crime, Report for the Home Office, University of Sheffield and University of East Anglia

Dolan, P. and Peasgood, T. (2007) Estimating the Economic and Social Costs of the Fear of Crime, British Journal of Criminology, **47**, 121-132

Donaldson et al: Weighting and Valuing Quality Adjusted Life Years: Report (2008)
http://www.pcpoh.bham.ac.uk/publichealth/methodology/docs/publications/JH12_Donaldson_Social_Value_of_QALY_July2008.pdf

Franic, D, Pathak, D, Gafni, A, (2003) Quality adjusted life-years was a poor predictor of women's willingness to pay in acute and chronic conditions: results of a survey, *Journal of Clinical Epidemiology,* 58 (3): 291-303

Gyrd-Hansen (2006) Willingness to pay for a QALY, *Health Economics*, 12, 1049 - 1060

Hammitt, J. K., & Liu, J. T. (2004) Effects of disease type and latency on the value of mortality risk. *Journal of Risk and Uncertainty,* **28**(1), 73-95

Horowitz, J. K., & Carson, R. T. (1993) Base-line risk and preference for reductions in risk-to-life. *Risk Analysis, 13*(4), 457-462

Johansson-Stenman, O., & Martinsson, P. (2008) Are some lives more valuable? An ethical preferences approach. *Journal of Health Economics, 27*(3), 739-752

Jones-Lee, M. W. and Loomes, G. (1995) Scale and context effects in the valuation of transport safety, *Journal of Risk and Uncertainty,* 11, 183-203

Kuppermann, M., Nease, R. F., Ackerson, L. M., Black, S. B., Shinefield, H. R. and Lieu, T. A. (2000) Parents' Preferences for Outcomes Associated With Childhood Vaccinations, *Pediatric Infectious Disease Journal,* **19**, 129-133

Lundberg et al, 1999 Quality of life, health state utilities and willingness to pay in patients with psoriasis and atopic eczema, *British Journal of Dermatology,* **141**, 1067-1075

Mason, H., Jones-Lee, M. and Donaldson, C. (2008) Modelling the Monetary Value of a QALY: A New Approach Based on UK Data, Health Economics online, DOI: 10.1002/hec.1416

McDaniels, T. L., Kamlet, M. S., & Fischer, G. W. (1992) Risk perception and the value of safety, *Risk Analysis,* **12**(4), 495-503

Mendeloff, J. M., & Kaplan, R. M. (1989) Are large differences in lifesaving costs justified - a psychometric study of the relative value placed on preventing deaths. *Risk Analysis,* **9**(3), 349-363

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

Pinto Prades, J., Loomes, G. and Brey, R. (2009) Trying to Estimate a Monetary Value for the QALY, forthcoming in *Journal of Health Economics*

Savage, I. (1993) An empirical-investigation into the effect of psychological perceptions on the willingness-to-pay to reduce risk. *Journal of Risk and Uncertainty,* **6**(1), 75-90

Slovic, P. (1987) Perception of risk. *Science*, **36**, 280-285

Subramanian, U. and Cropper, M. (2000) Public choices between life saving programs: The tradeoff between qualitative factors and lives saved. *Journal of Risk and Uncertainty*, **21**(1), 117-149

Sunstein, C. R. (1997) Bad deaths. *Journal of Risk and Uncertainty,* **14**, 259-282

Tsuge, T., Kishimoto, A., & Takeuchi, K. (2005) A choice experiment approach to the valuation of mortality. *Journal of Risk and Uncertainty,* **31**, 73-95

Van Houtven, G., Sullivan, M. B., & Dockins, C. (2008) Cancer premiums and latency effects: A risk tradeoff approach for valuing reductions in fatal cancer risks. *Journal of Risk and Uncertainty, 36*(2), 179-199.

Vassanadumrongdee, S. & Matsuoka, S. (2005) Risk perceptions and value of a statistical life for air pollution and traffic accidents: Evidence from Bangkok, Thailand. *Journal of Risk and Uncertainty, 30*, 261-287.

	First author (year)	Methods
i	Jones-Lee (1985)	1103 Ps (stratified random sample UK population) indicated preferences for saving 100 lives from either motor accidents (4,000), heart disease (11,000), cancer (16,000). Numbers in brackets show data presented about numbers of deaths each year in people aged under 55.
ii	Mendeloff (1989)	103 Ps (convenience samples of 50 psychology undergraduates/ 35 health policy students/ 18 retirees) assigned priorities to preventing 8 programs each of which prevented one fatality : limit workplace exposure to cancer-causing chemicals, bike lanes to protect cyclists, enforce regulation of crib slat width , remove roadside obstacles to protect motorists , increase paramedics/ ambulances to reach heart attack victims more quickly, enforce air pollution standards to benefit people with lung disease 10 years in future , build barriers on roads to prevent head-on collisions, enforce standards to prevent falls at construction sites . Ps assigned ratings to programs relative to a rating of 10 to their lowest ranked program.
iii	McDaniels (1992)	55 Ps (convenience sample of US adults) provided household WTP estimates for 20% proportional reductions in annual deaths from 5 well-defined hazards - with data provided on deaths per year (in brackets): automobiles (10,000), aviation (200), workplace chemical VCM (1), power tools (80), liquefied natural/ petroleum gas (2); and 5 less well-defined hazards – no data provided on deaths per year: chlorinated water, hazardous waste, nuclear energy, sulphur air pollution, electromagnetic fields. Hazards were also rated on 8 risk perception scales: voluntariness, severity, knowledge, control, dread, personal exposure, public exposure, overall risk.
iv	Savage (1993)	1027 Ps (random sample of adult population) divided up a $100 charitable donation to lower the risks of airplane accidents, household fires, automobile accidents and stomach cancer – and rated their level of dread (how nervous thinking about the hazard made them feel), knowledge (how informed they are about the risk) and personal threat.
v	Horowitz (1993)	In two surveys of 68 and 104 Ps (convenience sample) ranked 8 substances in terms of their preference for spending government funds on reducing risks (e.g., pesticides in drinking water, automobile exhaust). They also ranked the current and future number of deaths caused by each substance, personal risk, percentage of US exposed to the substance, effects on the environment and government effectiveness in preventing deaths.
vi	Cropper (1994)	3000+ Ps (telephone survey of general public) were asked to choose between a program that would save 200 lives from diseases and kill 20-year olds [or 40-year olds] and a program that would save [100, 250… 6,000] lives from diseases that kill 60-year olds.
vii	Jones-Lee (1995)	225 Ps (quota sample of London Underground users) rated their agreement/disagreement with statements comparing LU accidents with car accidents in terms of their level of dread, responsibility for their own safety, choice over exposure to risk, and control over own safety. Matching questions were used to elicit the number of fatalities in small-scale road accidents equivalent to 25-30 fatalities in small-scale Underground accidents.
iii	Baron (1997)	29 Ps (convenience sample of students) were asked how much money out of a budget of $100 billion they were allocate to reducing 2,600 deaths per year from a list of 18 causes (including homicides and car accidents). Ps also estimated the annual death rate (prevalence per 100,000).
ix	Sunstein (1997)	116 Ps (convenience sample of law students) were asked how many heart attack deaths and deaths from airplane crashes would equal 100 deaths saved by cancer
x	Cookson (2000)	52 Ps (parents from a local primary school) were asked their WTP in extra taxes over 10 years to prevent 100 deaths from car accidents, food poisoning, rail accidents, medical radiation, birth control pills and air pollution. and rail accidents. They were also asked to assign a relative monetary value to the prevention of 100 deaths from each context relative to a monetary value of £100million for car accidents. Matching questions were also used to relative values for preventing deaths from car accidents, food poisoning, and rail accidents.

	First author (year)	Methods
xi	Subramanian (2000)	1000 Ps (random sample in telephone survey) were asked to choose between pairs of programs with varying ratios of lives saved (e.g., smoking education vs. air pollution control, colon cancer screening vs. water pollution control) and rate the programs on qualitative characteristics of blame, controllability, seriousness, personal risk, and lag before program saves lives.
xii	Covey (2001)	147 Ps (general population quota sample) chose between hypothetical life-saving programs with different baseline numbers of deaths per year caused by the hazards (i.e., Hazard L causes 20-25 deaths per year, Hazard H causes 500-600 deaths per year)
xiii	Chilton (2002)	In 1998 130 Ps (quota sample of UK general population) and in 2000 150 Ps (quota sample of UK general population and regular rail users) provided matching responses for preventing deaths from road accidents, rail accidents, fires in public places and domestic fires. They also compared either domestic fires or rail accidents with road accidents on the qualitative characteristics of scale, personal-control, voluntariness, media-attention, expert-knowledge, uneasiness, deaths-per-year, age-groups-affected, personal-risk and household-benefit
xiv	Carlsson (2004)	996 Ps (random sample of Swedish citizens) provided WTP estimates for a safer air or taxi journey which reduced their fatality risk from 1 in a million to 0.5 in a million. The original prices of the journeys were varied (50 or 500 SEK for taxi, 500 or 3000 SEK for air).
xv	Hammitt (2004)	1248 Ps (random telephone survey of Taiwanese) provided WTP estimates for protecting their household from four environmental health risks which differed according to whether the disease is latent (symptoms occurs in 20 years after which 2-3 years life expectancy) or acute (symptoms begin in a few months with 2-3 year life expectancy), cancer or non-cancer, whether it affects the lung and is caused by contaminated water or the liver and is caused by industrial air pollution. Symptom descriptions for all scenarios identical.
xvi	Tsuge (2005)	400 Ps (random sample of Tokyo residents) undertook a choice experiment to estimate their private WTP for reductions in mortality risks either now or in the future from cancer, heart disease and accidents. Each risk was rated on voluntariness, controllability by government, dread(pain), dread(fear), severity, private knowledge, public knowledge, and personal exposure.
xvii	Vassanadum-rongdee (2005)	Ps (524 in air pollution study, 301 in traffic accidents – Bangkok residents) provided WTP estimates for reducing private mortality risks from air pollution (annual health check) and traffic accidents (airbags) and rated perceptions of voluntariness, severity, controllability, dread, personal exposure, public exposure, immediacy, personal knowledge and expert knowledge.
xviii	Covey (2006)	Ps (1033 in CAPI survey 1957 in Internet survey – both random samples of UK general population) provided matching responses for preventing different types of rail deaths (e.g., rail passenger killed in a collision accident, child trespasser killed taking shortcut across a track with inadequate fencing)
xix	Johansson-Stenman (2008)	1444 Ps (random sample of Swedish population) undertook a choice experiment between road investment projects that varied in terms of the number of saved individuals, their age-groups and whether they were car drivers or pedestrians.
xx	VanHoutven (2008)	788 Ps (online WebTV survey of US residents) chose between locations to live for a year which varied in terms of the car accident deaths and fatal cancers (stomach, liver or brain cancer) caused per year. The assumed latency periods for the cancer were varied from 5 to 25 years and the assumed morbidity period from 2 to 5 years.