

**A Review of Methodological Issues in the
Conduct of WTP Studies in Health Care II:
Administration of a CV Survey**

Mr Richard Smith

Senior Lecturer, Health Economics Unit, Centre for Health Program Evaluation
Monash University

Professor Jan Abel Olsen

Visiting Professor, University of Tromsø, Norway

Mr Anthony Harris

Senior Lecturer, Health Economics Unit, Centre for Health Program Evaluation
Monash University

CENTRE PROFILE

The Centre for Health Program Evaluation (CHPE) is a research and teaching organisation established in 1990 to:

- undertake academic and applied research into health programs, health systems and current policy issues;
- develop appropriate evaluation methodologies; and
- promote the teaching of health economics and health program evaluation, in order to increase the supply of trained specialists and to improve the level of understanding in the health community.

The Centre comprises two independent research units, the Health Economics Unit (HEU) which is part of the Faculty of Business and Economics at Monash University, and the Program Evaluation Unit (PEU) which is part of the Department of General Practice and Public Health at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

PUBLICATIONS

The views expressed in Centre publications are those of the author(s) and do not necessarily reflect the views of the Centre or its sponsors. Readers of publications are encouraged to contact the author(s) with comments, criticisms and suggestions.

A list of the Centre's papers is provided inside the back cover. Further information and copies of the papers may be obtained by contacting:

The Co-ordinator
Centre for Health Program Evaluation
PO Box 477
West Heidelberg Vic 3081, Australia
Telephone + 61 3 9496 4433/4434 **Facsimile** + 61 3 9496 4424
E-mail CHPE@BusEco.monash.edu.au

ACKNOWLEDGMENTS

The Health Economics Unit of the CHPE is supported by Monash University.

The Program Evaluation Unit of the CHPE is supported by The University of Melbourne.

Both units obtain supplementary funding through national competitive grants and contract research.

The research described in this paper is made possible through the support of these bodies.

AUTHOR(S) ACKNOWLEDGMENTS

This paper forms part of a literature review commissioned by the Australian Department of Health and Aged Care to promote informed discussion of the measurement and valuation of health outcomes. It falls within the consultative process of revising the November 1995 edition of The Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee: including major submissions involving economic analyses. The views expressed in the paper are not necessarily those of the Department, the Pharmaceutical Benefits Advisory Committee or its Sub-Committees and do not necessarily reflect the final outcome of this aspect of the next edition of these Guidelines.

ABSTRACT

Cost-effectiveness, and cost-utility, analyses have historically been the most widely used techniques of economic evaluation applied to the evaluation of health care programs. However, in recent years there has been renewed interest in the use of cost-benefit analysis, which requires the assessment of programme benefits in monetary terms. The emerging consensus is that such monetary valuation is most appropriately obtained using a survey of individual 'willingness-to-pay' (WTP) for the program of interest.

There are obviously a considerable number of methodological issues and potential biases to be considered in performing such a survey, which may be grouped into three main areas: (i) the construction and specification of the contingent market; (ii) the administration of the survey; and (iii) the analysis and interpretation of the WTP data. In addition, there are a few issues which also warrant consideration, such as assessing validity and reliability, and the impact of ability to pay and income distribution issues. This paper is concerned with assessing the administration of the CV survey. The other issues are considered in companion Working Papers.

Administration of the survey concerns the presentation of the contingent market to the respondent. The two main issues within this are the format of the questionnaire presented to the respondents, and the mode of administration of the questionnaire (critically, face-to-face interview or not). Other issues relevant within this broad area of presentation include, for example, the order of questions as posed to respondents, whether an explicit budget constraint is invoked, the sample from whom values should be elicited and possible bias caused by specifying values or by the interviewer. Each of these is discussed in this paper.

The paper considers the literature relating to the issues mentioned, and uses this to derive a set of 'recommendations' for current 'state of the art' conduct of WTP surveys with respect to survey administration. WTP studies conducted to date in health and health care are then reviewed with such recommendations in mind, to assess the degree to which they reflect this 'state of the art'. It is concluded that such studies perform poorly when judged in this manner.

TABLE OF CONTENTS

Abstract

1	Introduction	1
2	Literature Review Methodology	5
3	Questionnaire Format	8
4	Mode of Administration	13
5	Budget Constraint	15
6	Order Effects in Questions Posed	17
7	Sample Characteristics	19
8	First Response Bias	22
9	Compliance Bias	26
10	Protext Bids	28
11	Conduct Alongside Clinical Trials	30
12	Conclusions and Recommendations	31
	References	33
	Appendix 1: The complete Reference List of Papers Reviewed	39
	List of Tables	
Table 1:	Questionnaire Format	12
Table 2:	Pharmacoeconomic Relevance Qst Format	12
Table 3:	Data Collection Technique	14
Table 4:	Response Type by Implications	21

A Review of Methodological Issues in the Conduct of WTP Studies in Health Care II:

Administration of a CV Survey

1 Introduction

Cost-benefit analysis (CBA) has been the major practical result of welfare economic theory over the last 50 years (Olsen & Smith, 1999). It has been used and developed particularly in the fields of environmental and transport economics¹, and over the last decade interest in assessing the applicability of the technique to health care, and conducting CBAs of health care programs, has grown considerably².

Within CBA the benefits of an intervention are assessed in units commensurate with the cost, typically monetary units, and there are two main methods of eliciting these values. The first uses market information for complementary or substitute goods. This so called 'hedonic' or 'revealed preference' technique uses, for example, labour or housing markets to value risk and thus place a money value on the risk of death or injury (Viscusi, 1993). The second involves an experimental survey, or contingent valuation (CV)³, approach to individual monetary evaluation of hypothetical changes in health and welfare. It asks individuals about either the amount of compensation required (their willingness to accept compensation) or their willingness to pay (WTP) to avoid an illness (Drummond et al, 1997).

¹ Carson et al (1993), for example, list over 1,400 contingent valuation studies conducted in environmental economics since 1964.

² The first WTP study in health care is widely acknowledged to be one assessing the valuation of reductions in risk from mobile coronary care units (Acton, 1973). However, from then there were only another four studies conducted prior to 1990 (Thompson 1984, 1986; Berwick and Weinstein 1985; Reardon and Pathak 1989), with approximately 50 conducted since 1990.

³ So called because the answers to the valuation questions are contingent upon the particular hypothetical market described in the survey to respondents.

The second of these techniques, contingent valuation, has a strong basis in the theory of constrained utility maximization (Johansson, 1995) and has become the dominant technique used in the monetary valuation of benefits, in health care as well as other areas, in recent years; the main consensus being that individual 'willingness to pay' is the most appropriate means to value benefits in monetary terms (Olsen & Smith, 1999).

This paper is part of a series of papers undertaken to review the conceptual basis and methodological issues pertinent to the conduct of WTP studies in health and health care, and of pharmaceuticals in particular. The conceptual background, and its application in empirical studies, are discussed in two prior companion papers (Olsen et al, 1999; Olsen & Smith). The purpose of this paper is to review the methodological conduct of such studies. This includes an assessment of the current recommendations for "state of the art" contingent valuation studies, and an assessment of the manner in which CV studies in health care have been conducted to date, and how closely they match such recommendations.

Methodological issues are important at two levels. First, in the validity of the assessment made of the value of the benefits arising from a program. That is, does the WTP technique really measure what we think it does? Contingent valuation surveys (should) attempt to obtain precise and unbiased estimates of individual WTP for a program (Drummond et al, 1997). This means that CV surveys should be psychometrically robust (where apparently trivial changes in survey wording, information provided to the respondent, or question format, should not cause significant changes in the stated value), statistically reliable (stated values should be accurately estimated for the population, without significant bias), and economically sensible (stated values should correspond to other variables, such as income, in a manner predicted by economic theory).

Second, in order to use CBA, and therefore WTP, as a tool in decision making between various programs requires consistency in the use and reporting of analyses. Measurement should be driven by a coherent methodology to ensure we know what was valued, how it was valued, and the implications of changes in technique. It is therefore important that methodology in measurement is not left as an ad-hoc approach based on the disparate requirements of individual researchers. If methodology is not consistent then systematic biases⁴ will influence the result, giving an incorrect assessment of the benefits of a project.

Many such methodological issues and potential biases have been recognized and reviewed in a variety of publications over previous years, most notably in Mitchell & Carson (1989), Diamond & Hausman (1993) and Beattie et al (1997). However, such reviews have been either general with respect to the WTP technique, or specific to the environmental or transport sectors. This review will not attempt to repeat such detailed reviews of issues where unnecessary, but summarize the relevant findings of these previous papers to assess recommendations for the conduct of 'state of the art' surveys. The contribution of this review is to: (i) collate this relevant information into a succinct review of the major issues in WTP surveys; (ii) to discuss and consider the relevance

⁴ Biased responses are defined as those final WTP values that differ systematically from the 'true' values placed by respondents on the good in question.

and impact of these issues in the application of WTP to the evaluation of health care programs, especially the evaluation of pharmaceuticals; and (iii) to assess the conduct of WTP studies in health care to date with respect to the recommendations apparent concerning each of these issues/biases.

It is particularly important to assess the relevance/implication of these issues with respect to the health care sector. It is almost received wisdom in health economics that health care is a 'peculiar' commodity, and therefore any recommendations from other sectors should be considered in light of these 'peculiarities' (see for example Mooney (1986) and McGuire et al (1988)). For example, in the environmental literature it is thought that the questionnaire, to ensure 'realism and credibility' should resemble market conditions as closely as possible (for instance Brookshire et al, 1976). However, asking values in a market format in an area where consumers are not used to paying for the goods may give misleading responses and have different implications for the meaning of 'realism'. Similarly, although the NOAA Panel guidelines⁵ are appropriate for the use they were targeted, assessing non-use (or existence) values of environmental amenities, it is reasonable to question the rationality of **all** CV studies, including for use-value in health care, strictly adhering to them, no matter what the policy making context. Thus, although these issues have *general* relevance across whatever sector WTP is applied, there may be some *particular* implications emanating from use in the health care sector.

Although there are a considerable number of methodological issues and potential biases to be considered, it is possible to group these into three main areas of consideration when reviewing or conducting a WTP survey: (i) the construction and specification of the contingent market; (ii) the administration of the survey; and (iii) the analysis and interpretation of the WTP data. In addition, there are a few issues which also warrant consideration, such as assessing validity and reliability, and the impact of ability to pay and income distribution issues. This paper is concerned with the administration of the survey, the other two areas being considered in companion papers (Smith et al, 1999a; 1999b).

In a previous paper (Smith et al, 1999a), the authors considered the manner in which the contingent market is developed and defined. In some areas of that review the presentation of the market to respondents was briefly touched upon, such as when considering how to describe the commodity being considered in the scenario. This paper develops the theme of the presentation of the contingent market to the respondent further. The two main issues within this are the format of the questionnaire presented to the respondents, and the mode of administration of the questionnaire (critically, face-to-face interview or not). Other issues relevant within this broad area of presentation include, for example, the order of questions as posed to respondents, whether an explicit budget constraint is invoked, the sample from whom values should be elicited and possible bias caused by specifying values or by the interviewer. Each of these is discussed here, concentrating particularly on the first two issues mentioned.

⁵ The National Oceanic and Atmospheric Administration called a consensus panel together to review the state of CV studies for non-use value following the Exxon Valdez oil spill in Alaska (NOAA, 1993, 1994). A substantial set of guidelines were developed from this exhaustive review, which have been suggested for adoption in the conduct of most CV studies. However, they have not been without their critics (Hauseman, 1993).

Following this introduction, Section 2 describes the literature review process used in this paper. Subsequent sections then consider each of the issues outlined above. Assessment will be made of the issue, its characteristics, importance and remedial measures, by reference to published works, and a 'state of the art' approach discussed. Based on this 'state of the art' evidence, a recommendation concerning the handling of a particular issue/bias will be made. Within the discussion of each issue consideration will be made of the applicability and/or relevance of the issue/bias, and the recommendations concerning it, to the conduct of WTP in health care, specifically with respect to the evaluation of pharmaceutical products. Section 12 concludes with a subjective assessment by the authors on the 'validity' of CV surveys from a methodological viewpoint, and with a summary of recommendations concerning the review and conduct of CV studies of pharmaceuticals.

2 Literature Review Methodology

Within this paper are two literature 'reviews'. The first concerns the relevant literature, from health care and beyond, concerning each of the major methodological issues dealt with in this paper. These papers were not collected or reviewed as part of a 'systematic' review, but on a more ad-hoc basis. Many were already in the possession of one of the authors (RS) who already had a substantial collection of 'general' CBA/CV papers.

The second 'review' was of WTP studies specifically applied in the health care sector. The process for this is described here. Within each subsequent section of this paper the two reviews are presented as follows: first a review of the wider WTP literature relevant to the issue at hand, which will predominantly be from non-health care studies, followed by a brief summary of how studies conducted in health care have handled this issue, with an (obvious) indication of the quality of such studies.

2.1 Method for Selecting WTP Papers to Review

We have reviewed papers reporting from actual contingent valuation surveys of health or health care programmes published during the period from January 1st 1985 until May 31st 1998 (see Appendix 1). The selected papers were identified from three sources. First, a computerised bibliographic database search was conducted for papers written in English. Databases used for this search were Medline and EconLit, searched using the 'WinSpirs' package. The search was conducted using keywords (singularly and in combination with health and health care), as follows: contingent valuation, willingness to pay, willingness to accept, and cost-benefit analysis. Full details, including abstracts, were downloaded and reviewed for appropriateness and relevance to this review, with initial selection made by one of the authors (RS). This reduced list was also considered by the other author (JAO) and the final list used to order papers.

Second, during the review of these papers (once collected) any papers which looked to be of importance, and had been neglected by the above review, were noted and ordered. These additional papers were mostly focused upon specific issues within the method of WTP, CV or CBA, rather than empirical studies. In addition a search was made of a comprehensive in-house 'EndNote' database held by one of the authors (RS) concerning CBA and CUA.

Third, once the complete list of empirical studies was decided upon, this was sent to the four main authors in this area (Cam Donaldson, Magnus Johannesson, Bernie O'Brien and Mandy Ryan) asking them to identify: (i) whether all relevant papers of their own had been included; and (ii) whether they were aware of any other papers, not on the list, which they thought should be included in the review. The major contribution at this stage was to be forwarded copies of papers which were forthcoming, but at that present time (April 1998) had yet to have been published.

It should be noted that, as with any literature review, there is a possibility that literature may be overlooked due to: (i) being unpublished, or not published in peer-reviewed journals (such as consultancy reports for government or industry) and hence not on the on-line databases; (ii)

databases used not covering all relevant literature. We would appreciate to be made aware of any papers of relevance which we have not considered.

2.2 Exclusion Criteria

Papers had to report from a CV-survey of health care programmes or of dimensions of health. Four 'exclusion criteria' were applied: First, those which *did not report the results of a survey*. The majority of papers found in the review did not report from any specific WTP survey conducted, but addressed methodological or theoretical issues which could be of relevance to health. Second, papers which reported on the same survey as had been more extensively reported in other publications we had already included (Miedzybrodzka et al 1994, 1995; Johannesson 1992).

Third, papers which reported from surveys on WTP for the size of the health care sector (Eckerlund et al 1995) or the value of reduced waiting time (Johannesson et al 1998, Propper 1990). These papers made no reference to any health *outcomes*, nor to any health care *programmes*. Fourth, papers which were tangential to survey-based specific WTP, particularly those considering conjoint analysis (e.g. Ryan 1997). We found these to be inappropriate to this review as this method does not explicitly ask for a monetary value, but implies such values in a more indirect way.

2.3 Review Process

There were several specific criteria which were used by the authors to structure the review, relating to both conceptual and methodological issues. In this paper we shall be concerned with the following characteristics of the surveys:

- Data collection method (face-to-face, telephone, postal, self-administered)
- Development of scenario description
- Presentation of the scenario description (separate or as part of the question)
- Pharmacoeconomic relevance (explicit, implicit or no drug)
- Methodological and/or policy implications considered
- Payment vehicle used (out-of-pocket, tax, insurance)
- Types of respondents in the survey (eg users, general population)
- Cost-benefit statistic presented (eg NPV)
- Welfare measure used
- Time period of assessment and use of WTP (eg per year, per month)
- Questionnaire format (eg bidding, discrete)
- Order effects of questions
- Starting point/range bias tested for
- Interviewer bias tested for
- Strategic bias tested for
- Other biases assessed
- Duration of interview
- Summary statistic (eg mean, median)

-
- Use of confidence intervals
 - Type of statistical analysis used
 - Sample size
 - Response rate
 - Zero/high responses - frequency
 - Transformation of WTP values obtained
 - Income assessed
 - WTP adjusted for income effects
 - Presentation of income constraint to respondents
 - Validity and reliability assessment

The above criteria are used as the empirical basis within this paper. Both authors independently classified each study along these criteria, and then conferred to assess the degree of agreement. Areas upon which the authors had differed were then reviewed and a consensus decision made.

3 Questionnaire Format

Questionnaire format refers to the manner in which the WTP questions are presented, and consists of five basic forms: (i) open-ended questionnaires, which simply ask respondents for their maximum WTP; (ii) bidding questionnaire, which present respondents with an initial amount and they are bid up, or down, in increments (according to acceptance or rejection of the initial WTP value presented) until the respondents maximum WTP is achieved; (iii) payment cards (or categorical scales), where respondents are presented with a range of values and asked to indicate which they would pay (and sometimes which they would not); (iv) discrete questionnaires (also referred to as closed-ended, dichotomous-choice or binary), where respondents are presented with a single WTP value and asked if they would pay it or not; and (v) discrete with follow-up questionnaires, where the discrete question from (iv) is followed-up with a second (also called double-bounded discrete questions) and possibly a third (triple bounded) discrete question (Smith, 1997). Latterly the choice of format has come to be represented as a debate between those advocating the discrete (with or without follow-up) approach versus those who do not.

Historically, the favoured technique has been the ‘open-ended’ questionnaire, where respondents are asked directly for their maximum WTP for the commodity being valued. To combat poor response encountered with such direct questioning, some form of aid is usually used to assist the respondent (Johannesson, 1996; Jones-Lee, 1989). The most popular is the “bidding-game”, where a bid is made to the respondent, who then accepts or rejects it. The next bid offered is higher or lower, as appropriate, and this process continues until the respondents maximum WTP is reached. Another popular form of aid is the “payment card”, where the respondent is presented with a range of values from which to choose (Johannesson, 1996; Donaldson Shackely, Abdalla, 1997). However, there are three problems which have arguably been encountered in using this process: (i) low response rate (particularly for self-complete, mail based, surveys); (ii) strategic bias, where the respondent provides an artificially high or low bid to ‘game’ the process; and (iii) starting point (and range) bias (Johannesson, 1996; Mitchell & Carson, 1989). It is these three problems that primarily led to the development of the discrete choice approach.

In contrast to these other forms of question format, discrete choice questions offer respondents a single bid (value), which they either accept or reject. Because each respondent only provides this limited information (that their true maximum WTP is either equal, greater or less than the specified bid), the survey requires different sub-samples to be offered different bids. By varying the bid across sub-samples in this way it is possible to calculate the proportion of respondents who are willing to pay as a function of the bid. The resultant ‘curve’ is then interpreted as the aggregate demand curve for the commodity valued, if the proportion responding positively to the bid is multiplied by the number of respondents. The area under the curve is then taken as the societal WTP (Johannesson, 1996).

Three advantages are suggested for this approach: (i) that it more closely resembles consumer choice in an actual market situation, where price is given, which contributes to greater realism and respondent understanding, and therefore yields more ‘valid’ WTP values and higher response rates; (ii) that any one respondent cannot unduly influence the overall WTP value and thus provides less room for strategic behaviour (Mitchell & Carson, 1989; Hoehn & Randall,

1987); and (iii) it avoids starting point (and range) bias since respondents are only presented with one bid.

As with CV in general, this discrete choice method of eliciting WTP has been developed in the environmental economics field over the last decade, with gaining popularity (NOAA, 1993, 1994; Hanemann et al, 1991; Cameron & Quiggin, 1994; Hanemann, 1984; Alberini, 1995). A significant impetus to its wide scale adoption has been the recent influential recommendations of the United States National Oceanographic and Atmospheric Administration (NOAA) concerning the conduct of WTP studies for the valuation of environmental benefits (NOAA, 1994; Johannesson, 1996). A major recommendation of the NOAA panel was that discrete choice questions should be used in preference to 'open-ended' formats, concluding that responses to 'open-ended' questions were 'biased, erratic and unreliable'. The panel felt that the discrete format was more likely to lead to 'valid' WTP responses.

However, the discrete choice format does have potentially significant drawbacks. First, the information derived from each respondent is minimal, as their maximum WTP is not directly elicited and must therefore be inferred through statistical analysis. Since little information is gained from each respondent, discrete surveys require a sample size far in excess of that required for an open ended WTP survey (eg 1,000 to 2,500 (Kanninen, 1995; Alberini, 1995)), potentially involving greater research monies. Second, in order to infer WTP from discrete responses, some form of regression, or non-parametric statistical analysis, must be used. The most popular technique is logistic regression, where the probability of acceptance of a bid is calculated as a function of the bid (Johansson, 1995). Not only are fairly complex statistical procedures required, but also, in the case of parametric approaches, distributional assumptions are required which can greatly affect the resultant WTP estimate (Kristrom, 1990). Neither of these is necessary with analysis of WTP gained from other formats.

Thus, discrete choice surveys are highly statistically inefficient in that a very large number of observations is required to identify a distribution of values with any degree of accuracy (Alberini, 1995; Milon, 1989). An alternative strategy, intended to reduce this inefficiency (and therefore reduce the large sample sizes required), is to use a sequence of questions to narrow the range within which the respondents 'true' WTP is postulated to lie (Cameron & Quiggin, 1994; Hanemann, 1985; Carson et al, 1986). The respondent is presented with an initial bid, bid 1, which they either accept or reject. Following this, they are offered a second discrete bid dependent upon their response to the first. In general, if a respondent accepts the first bid the second bid choice will be offered at an amount approximately double the first, bid 2; if they reject this first bid then this second bid is set at approximately half the initial level, bid 3 (Hanemann et al, 1991). This sequence will thus yield the individuals value in one of four possible regions; zero to bid 3, bid 3 to bid 1, bid 1 to bid 2 or above bid 2 (where we may assume that bids will be bounded by the individuals income level). The result is a significant increase in the amount of information provided by each individual on the bounds within which their 'true' WTP is postulated to lie, which has been shown to be asymptotically more efficient than single bounded surveys (Hanemann et al, 1991). This approach is termed 'double bounding', with subsequent questions in the sequence yielding 'triple bounding' and so forth.

However, it is puzzling why the discrete technique has moved toward double or triple bounding. It is clear that there is some support for the use of these methods in obtaining more precise estimates of WTP and reducing the sample size required by a single bounded response. However, it would appear that to progress beyond a single bound creates a 'bidding' approach, with the number of bids simply truncated early in the process. The logic that argues for double or triple bounding would seem to apply equally well to n -bounding, which then brings us back to the 'traditional' bidding game. One might further argue that even the single bounded design is just a bidding game with one bid. It would seem, therefore, that there is some agreement that *some* form of bidding is appropriate. The question thus becomes one of how many bids, the relative benefits from using only one compared to more than one, and then the benefits of two versus three versus n . Double/triple bounding appears to offer the potential for bias of 'traditional' bidding, yet offer only limited data compared to that approach.

Despite these issues, it is fair to say that the discrete choice approach is now being advocated as the 'standard' format in questionnaire design, certainly within environmental economics. There has been little *critical* attention however to the use of such formats, especially with health care as compared to non-use values for environmental amenities, although one recent study (Donaldson et al, 1997) argued that biased results do not "necessarily lead to the conclusion of the NOAA Panel that questions be asked in a referendum format" (p154/5).

However, in a recent paper by one of the authors (Smith, 1997) the superiority of the closed-ended questionnaire was challenged. It was argued that whilst it is not *necessarily* the case that closed-ended questionnaires will resolve issues of bias, it certainly is the case that they result in two additional problems associated with statistical inference of the WTP values (or demand curve). It was concluded that

"It is undoubtedly the case that open-ended WTP surveys contain issues which need consideration, and the technique is far from being fully developed. However, to attempt to solve, or more appropriately avoid, these issues by adopting a very limited version of the traditional bidding game (using only 1, 2 or 3 bids) seems foolhardy. Not only might such a technique still present problems of survey cost and first-response bias, it also brings additional, unique, problems, such as issues concerning bid vectors and appropriate statistical analyses. These problems arise from the fact that, simply, discrete choice survey design and analysis is extremely complex This author would therefore suggest that the debate over discrete versus open-ended techniques is a 'red-herring' in the search for 'valid' WTP values derived from CV surveys. There appears no reason why the 'open-ended' survey should be summarily dismissed, and this author would recommend caution in the seemingly indiscreet adoption of the discrete choice question format at the present time." (Smith, 1997, pp22-23)

Perhaps as a final point, it is worth noting that, although estimates from open-ended and discrete surveys have been shown to provide markedly different WTP results, there is no conclusive theoretical or empirical justification for one being considered more 'valid' than the other. This is important to bear in mind when considering the current recommendations of most WTP researchers that discrete surveys are to be preferred as they are argued to be more 'valid'. ie that they may approximate market decisions more closely and hence provide more realistic data.

To resolve the issue there clearly needs to be empirical research which compares both types of survey with real choices. Until such studies are undertaken we are left with no definitive answer as to which is more 'valid', only a series of opinions from which the researcher must ultimately draw their own conclusion.

3.1 Recommendations

The basis for the use of discrete surveys in preference to open-ended seems to rest on some combination of: (i) a belief that this technique offers a more 'realistic' market, and will therefore lead to more valid responses (a truer estimate of actual WTP by respondents); and (ii) the supposed tendency of discrete surveys to yield higher response rates, through reduced mental demands (especially for mailed surveys). We would dispute both of these contentions, particularly as they apply in the health care context (although, of course, more research needs to be undertaken in this area as applied to health care commodities).

In the meantime, we would recommend against the use of discrete (single, double or triple bounded) surveys due to the vagaries of the technique enabling 'manipulation' of the final WTP result. For example, there is scope to 'bias' the resultant WTP result in two main areas: (i) the choice of bid vector, including bids offered respondents and the proportion offered those bids; and (ii) in the specification of functional form for parametric analyses. In contrast we would recommend that research is conducted to establish a 'common' format acceptable across WTP studies in health care. We would see this choice as being between: (i) payment cards; and (ii) some form of open-ended questionnaire with prompts, such as used by Jones-Lee (1985).

3.2 Health Care WTP Studies

We assessed five different forms of questionnaire: (i) open-ended; (ii) bidding games; (iii) payment cards; (iv) discrete questions; and (v) discrete with follow up (double or triple bounded) questions. There were also some combinations, generally where researchers were comparing such approaches. The results are presented in table 1.

Four studies did not allow an assessment of the questionnaire format. Of the remaining 50, 12 used payment cards, 12 used discrete surveys, seven bidding and seven open-ended. Only three used discrete and follow-up surveys. The three 'other' studies used a 'categorical' format, whereby the respondents fills in categories, which may be seen as approximate to payment card formats. In this case, the payment card is the most frequently used format, closely followed by discrete surveys. However, this result is influenced by the authors of studies, with seven of the discrete surveys undertaken by Johannesson in Sweden, and five of the payment cards by Donaldson and Ryan from Aberdeen.

Table 1: Questionnaire Format

Valid	Frequency	Percent	Cumulative percent
Open ended	7	13.0	13.0
Bidding	7	13.0	25.9
Payment card	12	22.2	48.1
Discrete	12	22.2	70.4
Discrete+FU	3	5.6	75.9
Other	3	5.6	81.5
Not available	4	7.4	88.9
Open+discrete	2	3.7	92.6
Open+paycard	1	1.9	94.4
Bidding+discrete	2	3.7	98.1
Open+other	1	1.9	100.0
TOTAL	54	100.0	

The relationship between questionnaire format and pharmacorelevance of studies is given in Table 2. The main feature is that no drug focussed study (explicit or implicit) used open-ended questionnaire, with discrete being the most popular. Again this is confounded by author, with five explicit, and five implicit, drug studies being conducted by Johanneson in Sweden.

Table 2: Pharmacoeconomic Relevance Qst Format

Pharmaco-economic relevance	QST FORMAT											TOTAL
	Open ended	Bid'g	Pay ment card	Dis-crete	Dis-crete +FU	Other	N/A	Open+dis-crete	Open +pay-card	Bid'g+dis-crete	Open + other	
Explicit drugs		2	1	3		1	1			1		9
Implicit drugs		2	3	3	3		1		1	1	1	15
Non-drug	7	3	8	6		2	2	2				30
TOTAL	7	7	12	12	3	3	4	2	1	2	1	54

4 Mode of Administration

Survey administration costs almost always comprise the largest element of a CV budget. Due to this they are often the target for cost-cutting measures in survey design, by the use of smaller samples or less expensive survey methodologies, such as mail-out questionnaires, or other forms of self-completion by the respondent. However, apart from the quality of the scenario and market context, the quality of the WTP values are only as good as the effort put in to obtaining those values (Mitchell and Carson, 1993).

Ideally, face-to-face interviews are to be recommended (Mitchell and Carson, 1993; NOAA, 1994). Such administration allows the presentation of a considerable amount of information whilst maintaining respondent interest and attention. It also allows control of the sequence of information presented, and also commands seriousness for the situation encouraging the respondent to carefully consider their response and take the matter as of importance (thus reducing the potential for 'hypothetical bias'). Indeed, "the superiority of in-person interviews is so large that a decision to use another mode of administration, such as telephone surveys, will usually need strong justification" (Mitchell and Carson, 1993, p28). For this reason the NOAA Panel recommended face-to-face administration (although with telephone administration as a possible second option in **extraneous** circumstances) (NOAA, 1994).

Mail surveys are a very poor third choice for administration. One of the most important reasons for this is the ability of people to read and comprehend the degree of information provided (Mitchell and Carson, 1989; NOAA, 1994). However, it is also the case that response rates are lower than interviews, which has implications for study results.

4.1 Recommendations

This issue is relatively uncontroversial - interviewer administered questionnaires are to be recommended, unless valid reasons why another technique is used can be provided. We would consider that arguing for mail-out administered questionnaires on the basis of cost is not sufficient reason (particularly given, in the case of pharmaceutical evaluations, the relative 'wealth' of pharmaceutical companies). This is a question of under resourcing of studies, and not of the appropriateness of the administration technique. It is possible that telephone interviews, if well-structured and conducted, can provide acceptable information, and it is recommended that this be seen as an acceptable 'second best' option provided a good case for why face-to-face interviewing is not possible.

4.2 Health Care WTP Studies

We reviewed studies according to whether data was collected by: (i) face-to-face interview; (ii) telephone interview; (iii) postal self-administration; (iv) other self-administration; or not available. The results are presented in table 3.

Table 3 **Data Collection Technique**

Valid	Frequency	Percent	Cumulative percent
Face2face	19	35.2	35.2
Telephone	6	11.1	46.3
Postal	12	22.2	68.5
Self adm	16	29.6	98.1
Not available	1	1.9	100
TOTAL	54	100.0	

The majority of studies, 34 (65%), did not undertake face-to-face interviews, and 28 (52%) undertook self-complete questionnaires. There did not seem to be any trend for mode of administration compared with questionnaire format, sample size or pharmacorelevance.

5 Budget Constraint

It may not occur to respondents that a WTP of a positive amount means they will have less available to spend on other goods and services. This leads to the potential for WTP to be over estimated. It is therefore important to ensure respondents realize the true 'opportunity cost' of their WTP in the sense of other goods and services they will have to forgo. It could also be argued that they should be made aware of other programs which may, or may soon, be available within the health care sector, otherwise there is a risk of allocating the whole of their health care budget to the one intervention before them (the 'embedding' effect) (Johansson, 1995).

The CV survey must therefore convey that provision of the commodity, and respondents' WTP, will have actual financial consequences for the household. Thus their budget constraint should be invoked, and respondents reminded of this. However, although it is clear that responses to WTP questions have to be made with a budget constraint in mind, it is not clear the extent to which this constraint, or the 'correct' constraint, should, or is, invoked in the decision making process. It is therefore argued that surveys should attempt to incorporate an explicit statement of budget constraint to the respondent to ensure that the appropriate constraint, household or personal income for example, is being invoked and used to constrain the WTP response.

Kemp and Maxwell (1993) argue that the most advanced method for doing this is the 'trade off analysis', a version of which is conjoint analysis (Green and Rao, 1971; Westwood et al, 1974). These authors argue that to avoid issues of 'noncommitment' or hypothetical bias, the survey should offer respondents a variety of commodities, or attributes of a commodity, to be considered within their fixed income, and be 'forced' to make trade offs between them. In this manner it is made clear to respondents the true, opportunity, cost of expressing a value for any one good, or attribute, in terms of the direct loss of consumption of the alternative goods or attributes. However, one should note that these techniques are not strictly CV surveys. They involve presenting respondents with a choice or *ranking* task in which they are asked to indicate preferences among several hypothetical options, each expressed in terms of its key attributes. They do not, in contrast with CV studies, require respondents to articulate an explicit value for any of the attributes or attribute differences under investigation. Rather, values are inferred indirectly from this preference data. However, as suggested, this technique has not been used a great deal in CV study, and Mitchell and Carson (1989) for example devote five out of 460 pages to this technique. Its use in health care has been advocated and used by Mandy Ryan of the Health Economics Research Unit (HERU) at Aberdeen University in Scotland, although it has not been taken up widely by other researchers.

A final point is that it is possible that the *perceived* budget constraint differs from the *actual*, or the one the researcher wishes to invoke (Mitchell and Carson, 1989). Thus, it is recommended, for example by the NOAA Panel (1994), that the respondent should be made **explicitly** aware of their specific budget constraint (income) or the one which the researcher wishes to invoke.

5.1 Recommendation

The respondent should be reminded of their budget constraint for their WTP, and if applicable given a specific budget constraint (depending upon type of study).

5.2 Health Care WTP Studies

We found that only six studies (11%) had included an explicit budget constraint, with little variation across time (the proportion of studies including a budget constraint remain comparable).

6 Order Effects in Questions Posed

It is possible that a sequence of questions might have an effect on a respondents WTP value (Mitchell and Carson, 1989). An example from the environmental literature illustrates this point. Samples and Holler (1990) asked respondents their WTP for preserving whales and seals. The sample was split into two: one receiving the question concerning the value of whales first, and then seals, the other the reverse order. Seal values were lower when the seal question was asked after the whale question, but the whale value was not affected by the sequence. The explanation offered was that “when respondents valued seals first, they used their behaviour in this market situation to guide their response to whale valuation questions. Since whales are generally more popular than seals, respondents were reluctant to behave more benevolently towards seals compared with humpback whales. Consequently, whale values were inflated in the S-W [seal then whale] questionnaire version to maintain a relatively higher value for the humpbacks. This behavioural anchoring effect did not exist in the W-S [Whale then seal] version, where whales were valued first” (P189).

Order effects mean that there is the potential for valuations to be made contingent upon the respondents ‘true’ value of the commodity *combined* with the valuation given, or portrayed, for another commodity. That is, valuations are not independent of prior valuations made. For example, if a respondent is asked for their WTP for a drug which relieves migraine 25% quicker than before, and then for one which relieves it 50% more, their valuation for the second may be higher than if they had not had the preceding valuation question, as they will relate their true value to that, perhaps concluding that because the ‘benefit’ is twice as great, their WTP value should be (which according to economic theory of diminishing marginal returns, would be higher than their ‘true’ value). Alternatively, respondents could provide a valuation to the initial valuation question corresponding to their total ‘account’ for ‘healthcare’, or ‘migraine’ expenditure. The second valuation question then presents them with a dilemma; they would not, in practice, pay any additional amount, but cannot face responding by saying that the value is equivalent as they may perceive that the interviewer will find them ‘stupid’ for doing so, and as a consequence they feel compelled to give a higher value. This has implications for ‘scope’ tests, as discussed in Smith et al (1999b). One possible solution could be to inform the respondent of all the valuation questions which will be asked initially, prior to any valuation, or enable them to go back and revise values once the exercise is complete.

6.1 Recommendations

If multiple valuation questions are to be asked then the survey should be constructed to provided a random ordering. In addition, analysis should place ‘question order’ as an independent variable in regression analysis of predictors of the WTP value to test for significance of this variable in determining (biasing) the WTP values.

6.2 Health Care WTP Studies

We classified studies according to, firstly, whether this was an issue or the study focussed on one commodity only, and for those which did consider more than one we classified them as using: (i) fixed order; (ii) random order; or (iii) other. The results where that for 41 studies (76%) this issue was not relevant, and for one we could not tell from the paper what was done. For the remaining 12 papers six used fixed orders and five random, with one using another means to address the issue of order effects. From this there appears to be some recognition that order effects are relevant and potentially important, although 50% of studies still are not taking account of this potentially significant bias.

7 Sample Characteristics

This has considerable overlap with the conceptual review of WTP (Olsen & Smith, 1999; Olsen et al, 1999). Olsen and Smith (1999) considered the issue of perhaps who should be included as a respondent within such a CV survey, and reviewed evidence concerning who had been included. The authors concern was that who was included as respondent would affect the types of values elicited (use value, option value, or 'caring externality' value): each of these three different values being activated according to whom the WTP is asked. These three values were considered to be differentially activated according to three broad types of respondents to a WTP survey; (i) currently diseased, (ii) currently non-diseased but at future risk, and (iii) currently non-diseased and not at future risk (O'Brien & Gafni, 1996). If current patients are asked, one would expect use value to be the predominant (although not exclusive) type of value expressed. That is the *ex post* perspective undertaken in those studies which have asked patients, or target groups with an identified need for the intervention in question. There are two different *ex ante* perspectives for subjects who could be at future risk.

The first is a private insurance premium which would express 'option' value. The nature of private insurance is to have the premium reflect one's own risk, hence there is no element of *ex ante* cross subsidization to express caring externalities. The second is a social insurance in the form of a tax contribution, which would express a combination of an option value for oneself plus a possible caring externality for others. Finally, subjects could be asked to express their WTP solely for other people's use of care. ie a caring externality. Only if a representative sample of the general population is asked would all motivations for WTP (use value, option value, and caring externality) be included⁶. However, the question of who should be asked becomes somewhat conceptual and ethical, dependent upon which of these sources of value one wishes to activate. If all sources of value are of concern then clearly the general population must be surveyed.

There is, however, a wider issue in the composition of the sample. Schelling (1968), has made an observation about the assumption of rationality in the models of WTP, that "however much 'rationality' we impute to our consumer, we must never forget is that the one thing he cannot control is his own imagination". This may mean that the individual has difficulty in controlling or understanding anxiety about death or about physical and psychological pain and may not be in a position to value them rationally. Thus the trade off of health care with other goods is difficult for the individual to make. To this end the importance and relevance of pursuing consumer sovereignty is questioned by those studies which do not do so, such as Donaldson (1990) who used relatives as agents for consumers, and those that do, such as Appel et al (1990) and Thompson et al (1984).

Further, although in the normal competitive market economists assume the knowledgeable consumer is the one whose preferences are to count, health care is considered different by many

⁶ It was noted that the vast majority of the surveys had asked users only, and that only nine (out of 54) studies had asked the general population (Olsen & Smith, 1999).

(Mooney, 1986), and as such we need to ask how relevant consumer sovereignty is to health care? Mooney (1986) suggests that we need to ask whether the individual is the best judge of their welfare, can they judge this welfare correctly, and do they want to make these decisions? He posits that the answers depend on the culture and health care system with respect to individual autonomy, that the problems of agency rule out individuals having knowledge of welfare, and that whether they want to make these decisions is an important ethical issue.

However, this debate is similar to that in the use of QALYs and we do not wish to become embroiled in the 'whose values' debate here. The main concern from a methodological point of view is whether the respondent can, in the case of health care interventions: (i) give a meaningful value, if they are presented with the fear and 'irrationality' that may be present when considering one's own health, particularly one's life; and (ii) can anyone other than those who have been through the treatment program really understand its attributes from the description given?

One final point is to raise an interesting, and potentially important, ethical point. Most CV studies within health care reviewed have been primarily *methodological* in focus. However, they have used *patients* for the sample values. It could be argued that this is unethical since testing a methodological issue does not depend (in the main) on the sample being patients or users of the service. They could, for example, be students and still produce usable results. However, to impose upon patients might be seen as unethical since they are taking part in a study which will not generate a present or future clinical benefit to them.

7.1 Recommendations

Whose values should be elicited presents a complex problem, but methodologically a simple resolution: it depends upon which primary values elicited by CV surveys one wishes to be activated. If it is narrow use-value then the users of the commodity may present the most tractable population, particularly as the description of the commodity will be backed up by personal experience of it. However, if in the evaluation of health care programs one wishes to activate wider values the sample will have to be correspondingly wider. An important point to note, however, is that if no decision is made and researchers are left to choose themselves, then the WTP values between products may differ solely because of the sample used, reflecting a difference in scope of values activated within the sample.

7.2 Health Care WTP Studies

From a conceptual point of view this issue was considered in the previous paper (Olsen & Smith, 1999). However, here we pick up the ethical concern raised of using patients for methodologically focussed studies. From Table 4, one can see that 15 studies were primarily aimed at methodological issues, yet focussed on users (predominantly patients).

Table 4: Response Type by Implication

RESPONSE TYPE	IMPLICATION			TOTAL
	Methodological	Health policy	Both	
Users	15	11	12	38
Conven sample	1	4	1	6
General pop	3	3	1	7
User+conven	1			1
User+genpop			1	1
Conven+genpop	1			1
TOTAL	21	18	15	54

8 First Response Bias

It is often argued that respondents find difficulty in locating their maximum WTP for a commodity, and instead try to establish a range, or indication, of its value using some other means. This phenomenon of 'anchoring' is a well known form of heuristic device in psychology, where absolute estimates are influenced up or down by the provision of different anchors; even if the respondent can see that the anchor provides no actual information concerning the 'correct' answer (Tversky & Kahneman, 1981; Tversky et al, 1988). The idea is that a respondent has largely uncertain views on the value of the product presented, and so forms WTP valuations based on prior experience, commodities which are 'close' and underlying beliefs. However, respondents, in wishing to have as much information as possible on which to base their value, may view the bid or prompt offered as representing what 'society' or 'experts' view as the value of the good, or what it may cost (Kahnemann et al, 1982, Donaldson et al, 1997). This is important for the bias it may introduce, but also because final estimates will be correlated with this anchor provided, and are thus open to manipulation by self-interested parties. The most common of these 'implied value cues' is 'first response bias', of which there are four common categories (Mitchell and Carson, 1989).

First, starting-point bias, where the elicitation method, or payment vehicle, (in)directly introduces a potential WTP amount that influences the WTP amount given by the respondents. In particular the concern that an initial prompt or bid in the bidding approach will anchor the respondent toward the starting bid, narrowing the distribution around the mean (portraying greater consensus than truly exists), and causing a loss in efficiency⁷.

The presence of starting-point bias in bidding games has been shown to be of concern (Rowe et al, 1980; Brookshire et al, 1980, 1981; Thayer, 1981; Boyle et al, 1985), including in some CV studies in health care (such as Stalhammer, 1997). For this reason the standard bidding game is not in widespread use at the present time. It is important to note, though, that there are instances where starting point bias has been tested for and not found (Kristrom, 1993, O'Brien and Viramonetes, 1994; O'Brien et al, 1998).

Second, range bias, where the elicitation procedure presents a range of potential WTP amounts which influences the WTP amount given by respondents. Payment cards are an obvious example of where this sort of bias may occur.

Third, relational bias, where the description of the good presents information to the respondent about its relationship to other public or private goods which influences the respondents WTP value.

Fourth, importance bias, where the act of being interviewed, or some feature of the instrument being used, suggests to the respondent that the good must have some value.

⁷ See Herriges & Shogren (1996) for an explanation of this.

It is argued that with discrete questions, since there is only one bid which is accepted or rejected, and these bids are varied across subsamples, there is no possibility of starting point or range bias (although these surveys are equally prone to relational and importance bias). With single bound questionnaires this may be the case, although it seems that the double/triple bounded discrete questionnaires would be susceptible to the same problem. Whilst efficiency may be improved, several authors (McFadden & Leonard, 1995; Hanemann et al, 1991) have found that the WTP estimates yielded by double bound are significantly different than those from single bound. There are various reasons this may occur (Hanemann et al, 1991; Cameron & Quiggin, 1994; Kanninen, 1995), but there is some evidence that starting point bias may be the cause (Herriges & Shogren, 1996).

This seems intuitive, since double, or triple, bounded surveys may be seen as merely a limited (and possibly more structured) version of the more traditional bidding approach to CV (Hanemann et al, 1991). The respondent, when answering the second discrete question, no longer compares their 'true' WTP to the follow up bid, but combines their 'true' WTP with the initial bid to form a revised WTP which is then compared to the follow up bid. Such bias has been found to be a potential factor in analysis (Herriges & Shogren, 1996). The implicit assumption of the double (triple) bound discrete approach is that the respondents answers to both (all) of the questions are driven by one underlying WTP value (Alberini, 1995; Milon, 1989). If this is true then the second (third) question increases information about the true WTP contained in the answer because it creates a tighter interval around this true WTP. However, if this is not true, and the two (three) values are definitely not identical (although highly correlated), then assuming they are can severely distort the estimated value distribution (Milon, 1989). The researcher is required, therefore, to explicitly recognize the endogeneity of the second discrete choice, thus providing support for possible starting point bias⁸ (Milon, 1989).

This bias may be in one of several directions. First, once a respondent has 'made a commitment' by indicating a willingness to pay for the first amount they are more likely to also be willing to pay the second amount than they would have had they not received the first offer. Second, respondents may interpret the first value as the average social value, or price, of the good and therefore not accept a higher bid, treating this as being asked to pay 'more than it costs'. Third, if respondents are not willing to pay the first amount they may feel guilty about it and be more willing to pay the second, lower amount, than they otherwise would. Finally, if they had said no to the first bid, they may become annoyed at being offered a lower bid, perceiving the interviewer to be trying to eke at least some money from them by lowering the bid, that they say no to the second bid even though they would have been likely to say yes had it been the only bid. Milon (1989), for example, concluded that the most dominant effect on the mean, and distribution of, WTP was that respondents who initially said yes were inclined to persist in saying yes, and those who initially say no are likely to persist in saying no. ie starting point bias.

⁸ It is quite possible that respondents have incomplete knowledge about their true valuation of a nonmarket resource and thus may give wrong yes/no answers on whether to pay or not. This logic is that if the respondent knew the value preference with certainty then they would have no difficulty in specifying it when facing an open-ended survey. However, given the experience that individuals face difficulty in doing this, it seems likely that there is some range within which the value lies, but as to the precise value the individual is uncertain. That is, the discrete format may help respondents who are 'information poor' regarding their actual WTP. However, this thesis applies equally to open-ended data, and so is not an issue here.

A close relation to starting point bias is 'yea-saying' bias, where the respondent will agree with amounts as offered by the interviewer. This form of bias is, particularly, of potential importance in discrete choice surveys (Hanemann et al, 1991; Holmes & Kramer, 1995). There is evidence to suggest that open-ended and discrete choice surveys do not demonstrate convergent validity (McFadden & Leonard, 1995; Kealy & Turner, 1993), and that this may be due to 'yea-saying', whereby respondents wish to agree with the interviewer regardless of their 'true' WTP. For example, the authors of one study comparing a payment card open-ended questionnaire and a single bound discrete questionnaire found that 'yea saying' had a positive influence on the results of the discrete responses compared to payment card approach (Holmes & Kramer, 1995). These authors conclude that "the results of this study strongly reject the hypothesis that dichotomous choice and payment card methods for the contingent valuation of a public good are procedurally invariant" (Holmes & Kramer, 1995, p130).

Another study, which specifically tested for issues of optimal bid design and potential bias in discrete choice (single and double bounded) surveys for CV, illustrated that 'yea saying' was a particular problem that might be encountered, with 20% of responses estimated to be 'yea-sayers' (Kanninen, 1995). For convergent validity, neither procedure is assumed to be a truer measure of the construct than the other. However, such studies are important as they empirically identify sources of bias in discrete surveys which might be hypothesized to be insignificant, such as 'yea-saying'.

A final point to be made is that even if starting point, or one of these other biases, does exist in any CV survey, open-ended or discrete, this may simply reflect a need by the respondents for some context, or 'bounds' to the question (Donaldson et al, 1997). The respondent may try to locate a 'price' of the good, or a close substitute or compliment (relational bias) to guide their valuation. For example, if I asked you for your maximum WTP for a new chocolate bar, your decision would be guided by the *price* of current chocolate bars, not your *maximum* WTP for it. If this is true, then we should not, necessarily, be looking to remove such bias, but finding means to assist the respondent in their valuation struggle. For example, by altering the 'context' of analyses to provide greater, or more detailed, information, such as suggesting the 'cost' of the intervention.

8.1 Recommendations

We believe that it is not possible to overcome some form of (implicit or explicit) first-response bias, simply because respondents will be looking for all the information they can on which to form their view of the 'worth' of a commodity to them. Within the pharmaceutical context in particular, there exists a 'pseudo-market' which may be guiding respondents views of 'price' or 'value'. We would therefore consider that this area of information should be researched, and recommend that bidding techniques not be used, with a standardization of questionnaires to open-ended with prompt or payment cards (as alluded to above).

8.2 Health Care WTP Studies

Starting point or range bias was only assessed in nine studies (17%). Of these, four had used the bidding approach, one the payment card and two discrete surveys (with one using combined open and payment cards and one combined bidding and discrete). None of these studies found it to be a significant factor.

9 Compliance Bias

Interviewer bias is the most widespread form of compliance bias. Here the respondent is either led, or wishes to, give a WTP amount which differs from their true WTP to either please the interviewer or because they are being led by the interviewer (cf 'yea-saying' bias). In using face-to-face, or telephone interviews, there is a risk that the valuation will be influenced (deliberately or accidentally) by the interviewer. Since CV surveys are often complex and more difficult to administer than 'standard' public polling or marketing surveys, they require consideration of professional trained interviews. For this reason Mitchell and Carson (1993) advocate professional interviewers, as they are specifically trained to adhere strictly to the wording of instruments, whereas others, such as graduate students or researchers themselves, may be tempted to adapt the survey in some fashion. Any sort of intervention, be it mendacious or well meaning, destroys the integrity of the survey, which is to obtain information on WTP from a sample of people who have been offered material which is *identical* and *consistent* between them. Any problems in presentation of material should be solved prior to the interview, not during it.

However, although some studies have pointed to possible interviewer bias, several studies which have undertaken to assess the effect of interviewers on WTP responses (through comparing interviewer and mail-out responses for example), have found no statistically significant difference in responses (eg. Schulze et al, 1983; Smith et al, 1993). It remains an ambiguous issue.

However, perhaps of more concern when considering evaluations of new programs undertaken by the manufacturer, such as with pharmaceuticals, is the other main form of compliance bias: sponsor bias. Here the respondent gives a WTP amount in an attempt to comply with the presumed expectation of the sponsor (or assumed sponsor) (Mitchell and Carson, 1989). In terms of pharmaceutical companies, for example, undertaking such surveys, sponsor bias could have a negative effect, in that respondents may react negatively to the sponsor (Johansson, 1995). Essentially, that WTP values will differ according to whether respondents are informed that the sponsor is a pharmaceutical company (such that the WTP amounts could simply be seen as contributing to the companies profits, and therefore stated WTP values will be lower than true WTP), or indeed according to which company is the sponsor. This latter possibility is important, as if there is a difference according which company then WTP values will differ systematically across CBAs based on which company the product is from, not the true value of it to the public. Arguably the former is not quite as significant an issue if all companies receive lower WTP valuations, since relativity between them should remain unaffected.

9.1 Recommendations

Since we would advocate the use of face-to-face (possibly telephone) interviews it is clear that there is the potential for interviewer bias. To guard against this, we would recommend that professional, trained and independent interviews are used in data collection. In addition, there should be an 'interviewer' variable within the regression analyses of factors which may explain the WTP values given. To guard against possible sponsor bias we would recommend that the sponsor of the survey is not named.

9.2 Health Care WTP Studies

The only form of compliance bias assessed was interviewer bias, although only by one study (out of the 19 who used face-to-face, and six who used telephone, interviews).

10 Protest Bids

In health state surveys, by definition the respondent would focus on health, while for WTP he/she will certainly also focus on having to sacrifice own income. The trade-off is fundamentally different; forego some money (representing utility potentially derived from consumption of other goods and services) in exchange for a health gain. Many people have aversion to this, either because they oppose the suggestion of paying more taxes, or they are against the idea of paying for health. Hence there is a serious problem of 'protest' answers (Cameron and Quiggin, 1994). These are typically zero bids (refusing to pay anything), although may be 'excessively' high (with the respondent protesting by offering a 'silly' amount). In order to guard against the 'influence' of protesters, compared with those expressing a 'true' valuation which is zero or high, it is important to treat all such bids with caution (Miller & Guria, 1991; Mitchell & Carson, 1989).

In recent years it has become practice for some researchers to ask those providing zero responses a series of questions to assess whether these are likely to be protest or 'true' values. This practice is to be recommended. Although such protest bids would generally be 'excluded' a high proportion of such bids could indicate a problem with the design of the CV survey, such as inappropriate bid vehicle. In this case it would seem appropriate to specify the proportion of such protest bids. However, although zero bids might be treated thus, there seems little symmetry in the treatment of 'high' bids, although *a priori* definition of 'high' may be difficult to establish.

10.1 Recommendations

Zero and very high bids to be reported and, if greater than 10% of responses, implications and possible reasons considered. In addition there should be a question in the survey to ask those who respond zero why they did so.

10.2 Health Care WTP Studies

We assessed the proportion of zero responses, and the proportion of respondents giving values more than twice the summary statistic (mean or median). For many studies we had to infer the values, where possible, as many did not explicitly report such values. We also therefore recorded whether the studies explicitly considered the possibility of protest answers resulting from the proportions of zero or high bids.

In 36 studies (67%) it was not possible to estimate the proportion of zero bids provided. For the remaining 18, seven recorded no zero responses, four reported 5% or less, two reported between 6% and 10%, two reported 10% to 30%, one 41%, one 65% and one 80%. There was no trend across the sample size used or questionnaire format. However, of five face-to-face interviews reporting zero responses, none was higher than 5%, with three having no zero responses, compared with two of the 10 self-complete questionnaires reporting no zero responses but three over 40% (including one at 80%). In terms of explicit drug studies, only one study could be assessed and was found to have no zero responses. Of four implicit drug studies, one had no zero bids, one had 1.5%, one 4% and one 14%.

In terms of high responses, data was only available for seven studies, of which one reported no high bids, two reported between 5% and 10%, two between 11% and 20%, and two between 21% and 30%. Again there was no discernable trend with sample size or questionnaire format, and only one face-to-face interview reporting a rate of 15%, and four self-administered reporting rates of 5% to 25%. However, the numbers are very small. Only 10 studies (19%) inquired into the possibility of protest bids, of which five had tested, or we could discern, the rates of zero and/or high bids.

11 Conduct Alongside Clinical Trials

For the evaluation of new treatments, such as pharmaceuticals, it may be desirable to conduct the CV survey alongside a clinical trial, as with other forms of evaluation. The issues concerned with conducting economic evaluations, principally applied to CEA and CUA, have been well discussed elsewhere, and many of these comments apply equally to the conduct of CBA (Drummond & Davies, 1991).

However, there may be some additional implications when applying a CV survey alongside clinical trials. For example, the choice of comparator. Will the study be a partial analysis of one versus nothing, as most CBA's in health care to date, or will it be compared with a placebo (and how might one therefore account for the 'placebo effect', or will it be compared with another close substitute (as in most of Donaldson's studies, for example (eg Donaldson 1990; Donaldson et al, 1997). If it is felt desirable that respondents have equal experience of a commodity, rather than only experience of one of the interventions (potentially leading to a bias in information available to make the valuation on), what are the implications for 'cross-over' trials? Would these contaminate the clinical trial? Of course there is also the implication that using those enrolled in a clinical trial will imply that only use-values will be assessed, as discussed in Olsen & Smith (1999).

11.1 Recommendations

We cannot see any reason why a CV study would need to be conducted alongside a clinical trial. The benefits of the intervention, such as in quality of life or recovery rates, cannot be adequately assessed until after the trial. Such outcomes should be used to develop the scenarios to be valued. Furthermore, a sample from a trial will (likely) consist solely of users, it is likely that the comparator will be a placebo (or the survey be partial) and there is not justification for gathering WTP values over time, as there is with other data. It is therefore our strong recommendation that such surveys not be conducted alongside clinical trials, but after, or at the least separately.

11.2 Health Care WTP Studies

It was difficult to assess this issue, although it was possible to assess whether the CV survey was conducted as part of a wider survey or as a specific event. Forty-two studies (78%) were specific, with the remaining 12 (22%) conducted as part of a wider survey or trial.

12 Conclusions and Recommendations

This review provides a comprehensive overview of the major issues involved in the administration of a CV survey. It represents the growing consensus, or lack of it, were appropriate, and reviews the conduct of WTP studies in health care to date. The review is necessarily subjective however, and there may be disagreements on the authors conclusions, interpretations and recommendations.

One caveat to note, however, is that many of the issues discussed here with respect to WTP should be seen in the wider context of trying to assess how individuals might behave in practice, where there is not an observable situation. Such problems associated with hypothetical surveys must affect all techniques whether WTP, QALYS or some other technique of assessing value. It is therefore important to try to distinguish between problems which are due to the inherent situation of trying to explain human behaviour in a hypothetical model, against those problems unique to the WTP method, compared to those specific to QALYS for example. No method is perfect so the choice becomes which level of imperfection one is content to live with.

Although it is difficult to provide rigid recommendations for how CV surveys should be administered, several key features do present themselves from our reading of the literature.

- 1 Questionnaire Format. We recommend that, at present, the discrete survey format is not advocated as standard practice, but rather the open-ended with prompt, or payment cards, is used. However, more 'standardization' of the range of prompts or values on a payment card needs consideration.
- 2 Mode of administration is perhaps one of the few areas of consensus, with a strong and universal recommendation that face-to-face interviews should be conducted unless there are **extraneous** circumstances for telephone surveys. Self-complete, particularly mail-based, surveys should not be used.
- 3 Order effects in questions posed should be acknowledged if multiple valuation questions are to be asked, and such a survey should be constructed to provide a random ordering. In addition, analysis should place 'question order' as an independent variable in regression analysis of predictors of the WTP value to test for significance of this variable in determining (biasing) the WTP values.
- 4 The characteristics of the sample of respondents is a key determinant of the WTP values provided, although this is not just a methodological issue (where we might suggest a random population sample for example), but is tied to the conceptual issues referred to in the previous paper concerning which aspects of value (use versus non-use) one might wish to activate.
- 5 Starting point, or range, bias might be present, especially given the format recommended above. However, all techniques to some extent are subject to this bias, and our recommendation to use pre-determined prompts, randomly allocated, will help reduce the impact of this, although we would recommend that the use or not of a prompt or range of

values be incorporated in regression analyses to determine whether its effect on the resultant WTP value is significant.

- 6 Compliance bias might be a problem - either in inflating, or deflating, the WTP values. We recommend that professional, and trained, independent interviews are used in data collection. We would also recommend that the sponsor of the survey, such as manufacture of the intervention, is not named.
- 7 Protest bids are possible, and we recommend that zero and very high bids are reported, and, if greater than 10% of responses, the implications and possible reasons for them considered. There should also be a question in the survey to ask those who respond zero why they did so.
- 8 The conduct of CV surveys alongside clinical trials is not recommended.

Reviewed against these recommendations, it is clear that WTP studies conducted to date have, overwhelmingly, performed poorly in the administration of CV surveys, although for many there is too little information to make any assessment. It is difficult, therefore, to assess the 'validity' of the WTP results, or to use the studies in a comparative sense to assist in priority-setting across program areas. There is an urgent need for a set of consistent guidelines concerning the administration of CV surveys to be debated and agreed upon if such studies are going to become a useful addition to the economic evaluation 'toolkit'.

REFERENCES

Alberini A 1995, 'Optimal designs for discrete choice contingent valuation surveys: Single bound, double bound, and bivariate models', *Journal of Environmental Economics and Management*, vol 28, pp 287-306.

Appel LJ, Steinberg EP, Powe NR, Anderson GF, Dwyer SA, Faden RR 1990, 'Risk reduction from low osmolality contrast media: What do patients think it is worth?'. *Medical Care*, vol 28, no 4, pp 324-337.

Beattie J, Chilton S, Cookson R, Covey J, Hopkins L, Jones-Lee M, Loomes G, Pidgeon N, Robinson A, Spencer A. 1997, Valuing health and safety controls: A literature review, Report to Health and Safety Executive, HM Govt. (Grant No. 3271/R73.04).

Berwick DM, Weinstein MC 1985, 'What do patients value? Willingness to pay for ultrasound in normal pregnancy', *Medical Care*, vol 23, no 7, pp 881-893.

Boyle K, Bishop R, Welsh M 1985, 'Starting point bias in contingent valuation surveys', *Land Economics*, vol 61 pp 188-194.

Brookshire DS, Ives BC, Schulze WD 1976, 'The valuation of aesthetic preferences', *Journal of Environmental Economics and Management*, vol 3, pp 325-346.

Brookshire DS, Randall A, Stoll JR 1980 'Valuing increments and decrements of natural resource service flows', *American Journal of Agricultural Economics*, vol 62, pp 478-488.

Brookshire DS, d'Arge RC, Schulze WD, Thayer MA 1981, 'Experiments in valuing public goods', in *Advances in Applied Microeconomics*, ed VK Smith, vol 1, Connecticut, JAI Press Inc.

Cameron T, Quiggin J 1994, 'Estimation using contingent valuation data from a 'dichotomous choice with follow-up' questionnaire', *Journal of Environmental Economics and Management*, vol 24, pp 218-234.

Carson TR, Hanemann WM, Mitchell RC 1986, Determining the demand for public goods by simulating referendums at different tax prices, Department of Economics Working Paper, University of California, San Diego.

Carson RT, Wright J, Alberini A, Flores N 1993, *A Bibliography of Contingent Valuation Studies and Papers*, ed CA La Jolla, Natural Resource Damage Assessment.

Diamond PA, Hausman JA 1993, 'On contingent valuation measurement of nonuse values', in *Contingent Valuation: A Critical Assessment*, ed JA Hausman, New York, North Holland.

Donaldson C 1990, 'Willingness to pay for publicly provided goods: A possible measure of benefit?', *Journal of Health Economics*, vol 9, no 1, pp 103-18.

Donaldson C, Shackley P, Abdalla M 1997, 'Using willingness to pay to value close substitutes: Carrier screening for cystic fibrosis revisited', *Health Economics*, vol 6, pp 145-159.

Donaldson C, Mapp T, Farrar S, Walker A, MacPhee S 1997, 'Assessing Community Values in Health Care: is the willingness to pay method feasible?', *Health Care Analysis*, vol 5, pp 7-29.

Drummond MF, Davies LM 1991, 'Economic analysis alongside clinical trials: revisiting the methodological issues', *International Journal of Technology Assessment in Health Care*, vol 7, pp 561-573.

Drummond MF, O'Brien B, Stoddart G, Torrance G 1997, *Methods for the Economic Evaluation of Health Care Programmes*, New York, Oxford University Press (2nd Edition).

Eckerlund I, Johannesson M, Johannsson P-O, Tambour M, Zethraeus N 1995, 'Value for money?: A contingent valuation study of the optimal size of the Swedish health care budget', *Health Policy*, vol 34, pp 135-143.

Green PE & Rao VR 1971, 'Conjoint measurement for quantifying judgemental data', *Journal of Marketing Research*, vol 8, pp 35-363.

Hanemann WM 1984, 'Welfare evaluations in contingent valuation experiments with discrete responses', *American Journal of Agricultural Economics*, August, pp 332-341.

Hanemann WM 1985, 'Some issues in continuous and discrete response contingent valuation studies', *Northeastern Journal of Agricultural Economics*, vol 5, pp 5-13.

Hanemann W, Loomis J, Kanninen B 1991, 'Statistical efficiency of double bounded dichotomous choice contingent valuation', *American Journal of Agricultural Economics*, vol 73, pp 1255-1263.

Hausman JA (Ed) 1993, *Contingent Valuation: A Critical Assessment*, New York, North Holland, 1993.

Herriges JA, Shogren JF 1996, 'Starting point bias in dichotomous choice', *Journal of Environmental Economics and Management*, vol 30, pp 112-131.

Hoehn JP, Randall A 1987, 'A satisfactory benefit cost indicator from contingent valuation', *Journal of Environmental Economics*, vol 14, pp 226-247.

Holmes TP, Kramer RA 1995, 'An independent sample test of yea-saying and starting point bias in dichotomous-choice contingent valuation', *Journal of Environmental Economics and Management*, vol 29, pp 121-132.

Johansson P-O 1995, *Evaluating Health Risks: An Economic Approach*, Cambridge University Press.

Johannesson M 1992, 'Economic evaluation of hypertension treatment', *International Journal of Technology Assessment in Health Care*, vol 8, pp 506-523.

Johannesson M 1996, *Theory and Methods of Economic Evaluation of Health Care*, Dordrecht, The Netherlands, Kluwer Academic Publishers.

Johannesson M, Johansson PO, Soderqvist T 1998, 'Time spent on waiting lists for medical care: An insurance approach', *Journal of Health Economics*, (forthcoming).

Jones-Lee MW 1989. *The Economics of Safety and Physical Risk*, Blackwell, Oxford.

Jones-Lee MW, Hammerton M, Philips PR 1985, 'The value of safety: Results of a national sample survey', *The Economic Journal*, vol 95, pp 49-72.

Kahnemann D, Slovic P, Tversky A 1982, *Judgement Under Uncertainty: Heuristics and Biases*, New York, Cambridge University Press.

Kanninen B 1995, 'Bias in discrete response contingent valuation', *Journal of Environmental Economics and Management*, vol 28, pp 114-125.

Kealy MJ, Turner RW 1993, 'A test of the equality of closed-ended and open-ended contingent valuations', *American Journal of Agricultural Economics*, vol 75, pp 321-331.

Kemp MA, Maxwell C. 1993, 'Exploring a budget context for contingent valuation estimates', in *Contingent Valuation: A Critical Assessment*, ed J.A. Hausman, New York, North Holland.

Kristrom B 1990, 'A non-parametric approach to the estimation of welfare measures in discrete response valuation studies', *Land Economics*, vol 66, pp 135-139.

-
- Kristrom B 1993, 'Comparing continuous and discrete contingent valuation questions', *Environmental and Resource Economics*, vol 3, pp 63-71.
- McFadden D, Leonard G 1995, 'Issues in the contingent valuation of environmental goods: Methodologies for data collection and analysis', in *Contingent Valuation: A Critical Assessment*, ed. JA Hausman, North Holland, Amsterdam.
- McGuire A, Henderson J, Mooney G 1988, *The economics of health care*, Routledge; London.
- Miedzybrodzka Z, Shackley P, Donaldson C, Abdalla M 1994, 'Counting the benefits of screening: A pilot study of willingness to pay for cystic fibrosis carrier screening', *Journal of Medical Screening*, vol 1, pp 82-83.
- Miedzybrodzka Z, Semper J, Shackley P, Abdalla M, Donaldson C 1995, 'Stepwise or couple antenatal carrier screening for cystic fibrosis? Women's preferences and willingness to pay', *Journal of Medical Genetics*, vol 32, no 4, pp 282-283.
- Miller T, Guria J 1991, *The Value of Statistical Life in New Zealand*, Land Transport Division, Ministry of Transport, Wellington, New Zealand.
- Milon JW 1989, 'Contingent valuation experiments for strategic behaviour', *Journal of Environmental Economics and Management*, vol 17, pp 293-308.
- Mitchell R, Carson R 1989, 'Using surveys to value public goods: The contingent valuation method', *Resources for the Future*, Washington, DC.
- Mitchell RC, Carson RT 1993, *Current issues in the design, administration and analysis of contingent valuation surveys*, Working Paper, Department of Economics, University of California, San Diego, Cal.
- Mooney GH 1986, *Economics of Medicine and Health Care*, Wheatsheaf, Brighton, Sussex.
- National Oceanic and Atmospheric Administration (NOAA), Arrow K, Solow R, Portney PR, Leamer EE, Radner R, Schuman H 1993, *Report of the NOAA panel of contingent valuation*, Federal Register 58, pp 4601-4614.
- National Oceanic and Atmospheric Administration (NOAA), Arrow K, Solow R, Portney PR, Leamer EE, Radner R, Schuman H 1994, *Natural resource damage assessments: Proposed rules*, Federal Register 59, pp 1062-1191.
- O'Brien B, Gafni A 1996, 'When do the 'dollars' make sense? Toward a conceptual framework for contingent valuation studies in health care', *Medical Decision-Making*, vol 16, pp 288-299.

O'Brien BJ, Viramontes JL 1994, 'Willingness to pay: a valid and reliable measure of health state preference?', *Medical Decision Making*, vol 14, pp 289-297.

O'Brien BJ, Goeree R, Gafni A, Torrance GW et al 1998 'Assessing the value of a new pharmaceutical: A feasibility study of contingent valuation in managed care', *Medical Care*, (in Press).

Olsen JA, Smith RD 1999. Who have been asked to value what? A review of 54 'willingness-to-pay' surveys in healthcare. *CHPE Working Paper No. 83*.

Olsen JA, Smith RD, Harris A 1999. Economic Theory and the Monetary Valuation of Health Care: an overview of the issues as applied to the economic evaluation of health care programs. *CHPE Working Paper No. 82*.

Propper C 1990, 'Contingent valuation of time spent on NHS waiting lists', *The Economic Journal*, vol 100, pp 193-199.

Reardon G, Pathak DS 1989, 'Assessment of a contingent valuation technique with utility estimation models', *Journal of Research in Pharmaceutical Economics*, vol 1, pp 68-89.

Rowe RD, d'Arge RC, Brookshire DS 1980, 'An experiment on the economic value of visibility', *Journal of Environmental Economics and Management*, vol 7, pp 1-19.

Ryan M 1997, 'Using conjoint analysis to assess women's preferences for miscarriage management', *Health Economics*, vol 6, pp 261-273.

Samples KC, Hollyer JR 1990, 'Contingent valuation of wildlife resources in the presence of substitutes and compliments', in *Economic Valuation of Natural Resources: Issues, Theory and Application*, eds RL Johnson, GV Johnson, CO Boulder, Westview Press, pp 177-192.

Schelling TC 1968, 'The life you save may be your own', in *Problems in Public Expenditure Analysis*, ed S. Chase, Washington (DC), Brookings Institution.

Sculze WD, d'Arge RC, Brookshire DS 1983, 'Valuing Environmental Commodities: Some Recent Experiments', *Land Economics*, vol LVIII; pp 151-72.

Smith RD 1997, *Contingent valuation: Indiscretion in the adoption of discrete choice question formats?*, Working paper 74, Centre for Health Program Evaluation.

Smith RD, Jan S, Shiell A 1993, *What are Australians willing to pay for road safety?*, Discussion Paper No. 21, Centre for Health Economics Research and Evaluation, University of Sydney, Sydney.

Smith RD, Olsen JA, Harris A 1999. A Review of Methodological Issues in the Conduct of WTP Studies in Health Care I: Construction and specification of the contingent market. *CHPE Working Paper No. 84*.

Smith RD, Olsen JA, Harris A 1999. A Review of Methodological Issues in the Conduct of WTP Studies in Health Care III: Issues in the Analysis and Interpretation of WTP data. *CHPE Working Paper No. 86*.

Stalhammar NO 1997, 'An empirical note on willingness to pay and starting point bias', *Medical Decision Making*, vol 16, pp 242-247.

Thayer MA 1981, 'Contingent valuation techniques for assessing environmental impacts: Further evidence', *Journal of Environmental Economics and Management*, vol 8, pp 27-44.

Thompson MS, Read JL, Liang M 1984, 'Feasibility of willingness to pay measurement in chronic arthritis', *Medical Decision Making*, vol 42, pp 195-215.

Thompson MS 1986, 'Willingness to pay and accept risks to cure chronic disease', *American Journal of Public Health*, vol 76, pp 392-396.

Tversky A. & Kahneman D 1981, 'The framing of decisions and the rationality of choice', *Science*, vol 453.

Tversky A, Sattath S, Slovic P 1988, 'Contingent weighting in judgment and choice', *Psychological Review*, vol 95, pp 371-384.

Viscusi WK 1993, 'The value of risks to life and health', *Journal of Economic Literature*, vol 1, no 4, pp 1912-19??.

Westwood R, Lunn T, Beazley D 1974, 'The trade off model and its extensions', *Journal of the Market Research Society*, vol 16, pp 222-241.

The Complete Reference List of Papers Reviewed

Appel LJ, Steinberg EP, Powe NR, Anderson GF, Dwyer SA, Faden RR 1990, 'Risk reduction from low osmolality contrast media: What do patients think it is worth?', *Medical Care*, vol 28, no 4, pp 324-337.

Bala MV, Wood LL, Zarkin GA et al 1998, 'Valuing outcomes in healthcare: A comparison of willingness to pay and quality-adjusted life years', *Journal of Clinical Epidemiology*, (forthcoming).

Berwick DM, Weinstein MC 1985, 'What do patients value? Willingness to pay for ultrasound in normal pregnancy', *Medical Care*, vol 23, no 7, pp 881-893.

Blumenschein K, Johannesson M 1998, 'The relationship between quality of life instruments, health state utilities and willingness to pay in patients with asthma', *Annals of Allergy, Asthma and Immunology* (in press).

Chestnut LG, Keller LR, Lambert WE, Rowe RD 1988, 'Measuring heart patients' willingness to pay for changes in angina symptoms', *Medical Decision Making*, 1996; vol 16, no 1, pp 65-77.

Culbertson VL, Arthur TG, Rhodes PJ, Rhodes RS 1988, 'Consumer preferences for verbal and written medication information', *Drug Intell Clin Pharm*, vol 22, no 5, pp 390-396.

Donaldson C 1990, 'Willingness to pay for publicly provided goods: A possible measure of benefit?', *Journal of Health Economics*, vol 9, no 1, pp 103-118.

-
- Donaldson C, Shackley P, Abdalla M, Miedzybrodzka Z 1995, 'Willingness to pay for antenatal carrier screening for cystic fibrosis', *Health Economics*, vol 4, no 6, pp 439-452.
- Donaldson C, Mapp T, Ryan M, Curtin K 1996, 'Estimating the economic benefits of avoiding food-borne risk: is 'willingness to pay' feasible?', *Epidemiol Infect*, vol 116, no 3, pp 285-294.
- Donaldson C, Shackley P 1997, 'Does "process utility" exist? A case study of willingness to pay for laparoscopic cholecystectomy', *Social Science & Medicine*, vol 44, no 5, pp 699-707.
- Donaldson C, Shackley P, Abdalla M 1997, 'Using willingness to pay to value close substitutes: carrier screening for Cystic Fibrosis revisited', *Health Economics*, vol 6, no 2, pp 145-159.
- Donaldson C, Thomas R, Torgerson DJ 1997, 'Validity of open-ended and payment scale approaches to eliciting willingness to pay. *Applied-Economics*, vol 29, no 1, pp 79-84.
- Donaldson C, Mapp T, Farrar S, Walker A, MacPhee S 1997, 'Assessing community values in health care: is the willingness to pay method feasible? *Health Care Analysis*, vol 5, pp 7-29.
- Donaldson C, Hundley V and Mapp T 1998, 'Willingness to pay: a method for measuring preferences for maternity care?', *Birth*, vol 25, pp 33-40.
- Eastaugh SR 1991, 'Valuation of the benefits of risk-free blood. Willingness to pay for hemoglobin solutions', *Int J Technol Assess Health Care*, vol 7, no 1, pp 51-57.
- Ellison P, Marr J 1994, 'Willingness to pay for unfunded health services in a family practice clinic', *Leadership in Health Services*, vol 3, pp 14-17.
- Fisher M, Marks A, Trieller K, Brody R 1985, 'Are adolescents able and willing to pay the fee for confidential health care?', *Journal of Pediatrics*, vol 170, pp 480-483.
- Gafni A, Feder A 1987, 'Willingness to pay in an equitable society: The case of the Kibbutz', *International Journal of Social Economics*, vol 14, no 1, pp 16-21.
- Gibb S, Donaldson C and Henshaw R 1998, 'Assessing strength of preference for abortion method using willingness to pay', *Journal of Advanced Nursing*, pp 30-36.
- Golan EH, Shechter M 1993, 'Contingent valuation of supplemental health care in Israel', *Medical Decision Making*, vol 13, pp 302-310.
- Gore PR, Madhavan S 1994, 'Consumers' preference and willingness to pay for pharmacist counselling for non-prescription medicines', *J Clin Pharm Ther*, vol 19, no 1, pp 17-25.
- Granberg M, Wikland M, Nilsson L, Hamberger L 1995, 'Couples' willingness to pay for IVF/ET', *Acta Obstet Gynecol Scand*, vol 74, no 3, pp 199-202.

Hassan EO, el-Nahal N, el-Hussein M 1994, 'Acceptability of the once-a-month injectable contraceptives Cyclofem and Mesigyna in Egypt', *Contraception*, vol 49, no 5, pp 469-488.

Johannesson M, Jonsson B, Borgquist L 1991, 'Willingness to Pay for Antihypertensive Therapy - Results of a Swedish Pilot Study', *Journal of Health Economics*, vol 10, no 4, pp 461-474.

Johannesson M, Aberg H, Agreus L, Borgquist L, Jonsson B 1991, 'Cost-benefit analysis of non-pharmacological treatment of hypertension', *J Intern Med*, vol 230, no 4, pp 307-312.

Johannesson M, Fagerberg B 1992, 'A health-economic comparison of diet and drug treatment in obese men with mild hypertension', *Journal of Hypertension*, vol 10, no 9, pp 1063-1070.

Johannesson M, et al 1993, 'Willingness to pay for lipid lowering: A health production function approach', *Applied-Economics*, vol 25, no 8, pp 1023-1031.

Johannesson M, et al 1993, 'Willingness to pay for Antihypertensive therapy -- further results', *Journal of Health Economics*, vol 12, no 1, pp 95-108.

Johannesson M, Johansson PO 1997, 'Quality of life and the WTP for an increased life expectancy at an advanced age', *Journal of Public Economics*, vol 65, no 2. pp 219-228.

Johannesson M, O'Connor RM, Kobelt-Nguyen G, Mattiason A 1997, 'Willingness to pay for reduced incontinence symptoms', *British Journal of Urology*, vol 80, pp 557-562.

Kartman B, Andersson F, Johannesson M 1996, 'Willingness to pay for reductions in angina pectoris attacks', *Medical Decision Making*, vol 16, pp 248-253.

Kartman B, Stalhammar NO, Johannesson M 1996, 'Valuation of health changes with the contingent valuation method: a test of scope and question order effects', *Health Economics*, vol 5, no 6, pp 531-541.

Krupnick AJ, Cropper ML 1992, 'The effect of information on health risk valuations', *Journal of Risk and Uncertainty*, vol 5, pp 29-48.

Lee SJ, Neumann PJ, Churchill WH, Malachowski ME, Weinstein MC, Johannesson M 1997, 'Patients' willingness to pay for autologous blood donation', *Health Policy*, vol 40, pp 1-12.

Lee SJ, Liljas B, Neumann PJ, Weinstein MC, Johannesson M, 'The impact of risk information on patients' willingness to pay for autologous blood donation', *Medical Care* (in press).

Lindholm L, Rosen M, Hellsten G 1994, 'Are people willing to pay for a community-based preventive program', *Int J Technol Assess Health Care*, vol 10, no 2, pp 317-324.

Nathan RG, Bont GM, Minz RB 1994, 'Patient interest in receiving audiotapes of information presented by their physicians', *Arch Fam Med*, vol 3, pp 509-513.

Neumann PJ, Johannesson M 1994, 'The willingness to pay for in vitro fertilization: A pilot study using contingent valuation', *Medical Care*, vol 32, no 7, pp 686-699.

O'Brien BJ, et al 1995, 'Assessing the economic value of a new antidepressant: A willingness-to-pay approach', *Pharmacoeconomics*, vol 8, no 1, pp 34-45.

O'Brien BJ, Goeree R, Gafni A, Torrance GW et al 1997, 'Assessing the value of a new pharmaceutical: A feasibility study of contingent valuation in managed care', *Medical Care*, (forthcoming).

Olsen JA, Donaldson C 1998, 'Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programs', *Social Science and Medicine*, vol 46, no 1, pp 1-12.

Osmond MH, Klassen TP, Quinn JV 1995, 'Economic comparison of a tissue adhesive and suturing in the repair of pediatric facial lacerations', *Journal of Pediatrics*, vol 126, pp 892-895.

Pennie RA, O'Connor AM, Garvock MJ, Drake ER 1991, 'Factors influencing the acceptance of Hepatitis B vaccine by students in health disciplines in Ottawa', *Canadian Journal of Public Health*, vol 82, pp 12-15.

Phillips K, et al 1997, 'Willingness to pay for Poison Control Centres', *Journal of Health Economics*, vol 16, no 3, pp 343-357.

Ramsey SD, Sullivan SD, Psaty BM, Patrick DL 1997, 'Willingness to pay for antihypertensive care: evidence from a staff-model HMO', *Social Science and Medicine*, vol 44, no 12, pp 1911-1917.

Reutzel TJ, Furmaga E 1993, 'Willingness to pay for pharmacist services in a Veterans Administration Hospital', *Journal of Research in Pharmaceutical Economics*, vol 5, no 2, pp 89-114.

Ryan M 1996, 'Using willingness to pay to assess the benefits of assisted reproductive techniques', *Health Economics*, vol 5, no 6, pp 543-558.

Ryan M 1997, 'Should government fund assisted reproductive techniques? A Study Using willingness to pay', *Applied Economics*, vol 29, no 7, pp 841-849.

Ryan M, Ratcliffe J, Tucker J 1997, 'Using willingness to pay to value alternative models of antenatal care', *Social Science & Medicine*, vol 44, no 3, pp 371-380.

Slothuus U 1997, Comparison of two contingent valuation methods: closed-ended questions with and with-out follow-up. Paper presented at the 18th Nordic Health Economists' Study Group Meeting, August 1997, Finland.

Stalhammer NO 1996, 'An empirical note on willingness to pay and starting-point bias', *Medical Decision Making*, vol 16, pp 242-247.

Thompson MS 1986, 'Willingness to pay and accept risks to cure chronic disease', *Am J Public Health*, vol 76, no 4, pp 392-396.

Weaver M, Ndamobissi R, Karnfield R, et al 1996, 'Willingness to pay for child survival: Results of a national survey in Central Africian Republic', *Social Science and Medicine*, vol 43, pp 985-998.

Zethraeus N 1998, 'Willingness to pay for hormone replacement therapy', *Health Economics*, vol 7, pp 31-38.