

**A Review of Methodological Issues in
the Conduct of Willingness-to-Pay Studies
in Health Care I:**

**Construction and Specification of the
Contingent Market**

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 Economic 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 construction and specification of the contingent market. The other two areas being considered in companion Working Papers.

This contingent market, fundamentally, consists of four elements: (i) a description of the (benefits of the) commodity to be valued (the 'scenario'); (ii) the means by which the WTP value will be administered (payment vehicle); (iii) whether the commodity is valued under conditions of certainty or risk; and (iv) the time period for valuation. Each of these elements of the contingent market is dealt with in this paper. In addition, there are two important biases which may occur as a result of the contingent market specification: (i) hypothetical bias, where WTP responses to the exercise do not represent actual valuations; and (ii) strategic bias, where there are elements within the exercise which encourage 'gaming' of the exercise. Again, these two are 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 market construction. 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	Scenario Specification	8
4	Payment Vehicle	16
5	Certainty Versus Risk	19
6	Time Period for the WTP Estimate	22
7	Hypothetical Bias	25
8	Strategic Bias	28
9	Conclusions and Recommendations	30

References	32
-------------------	-----------

Appendix 1	38
-------------------	-----------

List of Tables

Table 1	Scenario Development	12
Table 2	Scenario Development by Response Type	13
Table 3	Scenario Presentation	14
Table 4	Scenario Presentation by Pharmacoeconomic Relevance	14
Table 5	Scenario Presentation by Publication Period	15
Table 6	Payment Vehicle Used	17
Table 7	Payment Vehicle by Pharmacoeconomic Relevance	18
Table 8	Risk/Problems in the Scenario by Respective Type	21
Table 9	Time Period for WTP Question	23
Table 10	Time WTP by Pharmacoeconomic Relevance	24

A Review of Methodological Issues in the Conduct of Willingness-to-Pay Studies in Health Care I

Construction and Specification of the Contingent Market

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 maximisation (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, 1999). 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 and impact of these issues in the application of WTP to the evaluation of health care programs,

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.

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 those 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 assessing the construction and specification of the contingent market. The other two areas being considered in companion papers (Smith et al, 1999a; 1999b).

The construction and specification of the contingent market is dealt with in some detail as it is considered, by these authors at least, to be the most fundamental methodological issue in the conduct of WTP surveys. Contingent valuation requires respondents within a survey to consider the contingency of an actual market existing for a commodity and reveal the maximum they would be WTP to purchase that commodity (whether that be for personal 'use', as an option for future use, or for non-use reasons). Thus, fundamental to the WTP values obtained in this survey is the way in which the commodity is represented to respondents: "the central problem in a [contingent valuation] study is to make the scenario sufficiently understandable, plausible and meaningful to respondents" (p73, Johansson, 1995). That is, how the contingent market for the commodity is constructed and elements of it specified.

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).

This contingent market, fundamentally, consists of four elements: (i) a description of the (benefits of the) commodity to be valued (the ‘scenario’); (ii) the means by which the WTP value will be administered (payment vehicle); (iii) whether the commodity is valued under conditions of certainty or risk; and (iv) the time period for valuation. Each of these elements of the contingent market is dealt with in this paper. In addition, there are two important biases which may occur as a result of the contingent market specification: (i) hypothetical bias, where WTP responses to the exercise do not represent actual valuations; and (ii) strategic bias, where there are elements within the exercise which encourage ‘gaming’ of the exercise. Again, these two will be discussed here.

Following this introduction, Section 2 describes the literature review process used in this paper. Subsequent sections then consider each of the six issues outlined above in turn. 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 9 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. Two of the authors (RS & JAO) 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 Scenario Specification

“Only questions that create a realistic market for a precisely defined good can measure the type of income-constrained behavioural intention information suitable for use in economic analysis. The description of the good and the market context together comprise the conditions on which the willingness to pay decision is contingent”

Mitchell & Carson (1993, p10)

In designing the contingent market there are two primary goals: (i) to ensure that respondents clearly understand the characteristics of the commodity to be valued and the context in which it is offered; and (ii) to ensure that respondents find this market situation plausible. Both of these goals critically rely on the description of the market and commodity offered: the scenario.

The scenario, as it describes exactly what the respondent is to value, and (crucially) what we assume they are valuing, is the foundation of the resultant WTP value derived. It determines what the respondent feels they are getting for their proffered WTP. To illustrate, consider two surveys: survey A asks respondents the maximum they would pay for “an ultrasound”. Here the commodity is described in the vaguest of terms, and the impression of what the commodity involves and the benefits from it are left to the respondent to infer and decide, and the researcher is unaware of the actual benefit being valued. In addition, there is no indication of, for example, the payment vehicle or time frame involved (these, and other, issues are discussed later). Survey B, on the other hand might ask respondents if they would be WTP \$50 as a one-off out-of-pocket payment to secure a specified reduced level of specified side-effects from low osmolality contrast media, compared with the alternative of high osmolality contrast media currently used in radiography, with all other factors, such as risk, health status and information obtained, being equal between the two programs. This survey is preceded by interview material presenting information on the two programs and the purpose of the study. With a survey of type A we are essentially witnessing an *attitudinal* survey, of the sort used in opinion polls, to assess general attitude (good-to-bad) toward a program, or type of program (Mitchell and Carson 1993). The WTP aspect lacks any validity as there is no realistic information about either the program or the market context in which the valuation is given - these are left to the respondents imagination (risking hypothetical bias also (see later)): there is no clearly defined good or the necessary market context for a meaningful valuation. Survey B, by contrast, is more of a behavioural question, measuring an intention to pay for a *specific* program, with *specific* attributes, in a clearly defined fashion.

To predict successfully requires that the scenario offered the respondent be congruent with the actual decision context the researcher will be using the values for (Ajzen and Fishbein 1977; Kahneman et al, 1993). This means the features of the program must be clearly described, and presented in a clearly defined market. Respondents need to know what they are buying, who is going to provide it, how it will be provided, when it will be provided and how the respondent is expected to pay for it.

The importance of the scenario cannot be overemphasised, as it is the foundation to the WTP values obtained. It is hypothesised in economic theory that if the characteristics of a good change (for example more of the good is offered) then the WTP value for it will also change (for instance increase), to reflect the change in value, compared to other commodities available, brought about by the commodity's altered attributes. It is important to assess which factors are of importance (through pilot surveys) and to be sure to include precise specification of them in the survey. This is for two reasons. First, whenever scenarios fail to provide information about factors of importance to respondents, then respondents will fill these 'information gaps' with default assumptions (ie guesses) concerning the likelihood of the program working, what happens if the program does not work, what substitutes might be available and how long or often they might have to pay for the program (Fischhoff and Furby 1988). Such interpretation about key issues in the program being valued "render a CV study's estimates uninterpretable because there is no way of knowing what they were buying; even if this was known, it would differ from one respondent to another" (Mitchell & Carson 1993, p13). Second, if the program is described vaguely it is to be expected that WTP responses are vague, and thus, for example, insensitive to the scope of the program offered (this is discussed more later).

3.1 Understandability⁶

However, simply because the researcher *thinks* they have correctly specified the scenario does not mean that the respondent *understands* it as the researcher supposes: the scenario must be clearly communicated to the respondent (Mitchell and Carson 1993). Common mistakes include: (i) use of scientific words, or technical terms, which may be understandable by the researchers, but not understood, or misunderstood, by the respondent; (ii) cognitive overload in presenting too much detail; and (iii) insensitivity to leading wording or phrasing⁷. It is well known that even small changes in the wording of scenarios can cause large changes in the valuation of those scenarios (eg Schoemaker 1982; Tversky & Kahneman, 1981; McNeil et al, 1982; Gerard et al 1992; Smith & Dobson, 1993; Jones-lee, 1989).

Tversky and Kahneman (1981) for example have described many such "framing and labelling" effects, in which apparently 'irrelevant' changes in wording of the question (from the perspective of normative theory) produce substantial shifts in response. Their most famous example is one in which respondents choose between a risky and safe program to combat disease effects on the US population. The 'irrelevant' wording change is to describe the effects either in terms of lives lost or lives saved - although the objective outcomes are the same in each case. Respondents presented with losses make many more risk taking choices than those presented with gains (lives saved).

The essential problem with scenario development, in terms of framing and labelling effects, is summed by Schoemaker (1982), who stated that "problem representation [scenario] is inherently

⁶ Understandability is only half the issue, however; the scenario must also be believable if it is to be taken seriously and elicit non-arbitrary values. This issue is dealt with in the section on hypothetical bias.

⁷ It has also been suggested that pictures or diagrams best enable the good to be understood, rather than words (Appel et al, 1990; Miller & Guria, 1991; Thompson et al, 1984).

a subjective matter, [therefore] it is subject to only limited normative evaluation. Indeed, there exists no general normative theory as to how problems should be defined, or how language and context should be encoded” (p556). Due to these problems there is some argument that only those who have experienced the good should be assessed, but Mitchell and Carson (1993) make the important point that “prior experience with a good is not required, as some have argued, for respondents to understand its characteristics. While it is often easier to describe a good to a respondent who has had experience with it, this is not always the case.” (p18). Of course it may also not be possible to find sufficient people who have experienced the good, especially with a new intervention such as a pharmaceutical therapy. The question of course then becomes whether we can definitively state that one ‘framing’ of a problem is ‘better’ than another. However, at the very least the likely importance of such issues in each specific case can be assessed during testing of the survey.

Understandability is also a function of the context surrounding the program. The NOAA Panel recognise this, recommending a “high standard of richness in context to achieve a realistic background” (NOAA, 1993, p4608). Context can deal with any variable, but one of particular relevance would be the availability of close substitutes, and/or a description of what would happen in a state of the world without the program.

Having said that, the degree of information required to adequately describe the program, its context and the market within which it is offered may impose a significant cognitive burden on respondents. For this reason the researcher must endeavour to minimise the information which is needed, and communicate it in a clear and concise manner. The first can be achieved with extensive pre-survey assessment of relevant factors. The second can be achieved through following several techniques which have been used successfully. For example, Mitchell and Carson (1993) outline several procedures they used in conducting a CBA of the Exxon Valdez oil Spill, including:

- use of simple language, elimination of unnecessary words, and logical sequence of presentation of material;
- interspersing the narrative with questions to involve the respondent;
- using face-to-face interviews;
- careful training of interviewers to read material at appropriate pace, use of conversational inflection, pauses and eye contact;
- use of various visual display items.

It is, however, a sobering thought that even following these procedures the scenario they used still ran to 12 pages!

3.2 Potential Biases

Mitchell and Carson (1989), in what is widely considered to be the definitive work on methodological issues in conducting CV studies, focus upon the scenario in particular in their assessment of the methodology of CV surveys, and treat it very much as the focus of methodology - if the scenario is not constructed 'correctly' then the validity of the measure is in question regardless of subsequent measurement issues. The result is termed "scenario misspecification", and occurs when the respondent does not respond to the correct scenario (ie the one the researcher thinks they are responding too). There are three issues that Mitchell and Carson (1989) see as important in the scenario specification.

First, *theoretical* misspecification, where the scenario specified by the researchers is in fact incorrect, either in terms of economic theory or major policy relevance. This is misspecification that is caused by the researcher and should therefore be avoidable by correct specification of the question and the contingent market.

Second is *amenity* misspecification, where the perceived good being valued differs from that intended. This may occur in four ways: (i) where the respondent values a *symbolic* entity rather than the specific good (see 'embedding'); (ii) where the respondent values a larger or smaller entity than that offered (due for example to valuing spatial attributes of difference, or values part of the good, or more than the whole); (iii) where the respondent values the commodity on a different metric scale than intended; and (iv) where the respondent values a commodity containing *probability*, and misperceives the probability.

Third, the perceived *context* may differ from that intended. This involves a whole range of variables, some of which are considered in other sections, such as: (i) payment vehicle; (ii) property rights; (iii) method of provision; (iv) budget constraints; (v) the elicitation question itself failing to convey that the maximum WTP is to be given; and (vi) question order, where the sequence of questions can have an effect on WTP values given.

3.3 Recommendations

From the literature reviewed, a number of suggestions for designing and specifying scenarios may be advanced.

Of utmost importance is for the researcher to specify the characteristics, or attributes, of the commodity providing benefit. That is, it should be clear what the benefits of the commodity are. Invaluable here will be focus groups, to consider the views of differing groups, such as experts, general population and users of the commodity.

The characteristics, or types of benefit, should match the policy question and the underlying theoretical construct. For example, in the case of societal benefits, the scenario should include all elements of use, option and externality value. For a more narrow health care viewpoint, it may be

necessary to split out such values, and perhaps concentrate on assessing the value of the health effects alone, or values for use, option and externality value separately.

The commodity, these benefits and other elements of the contingent market (such as payment vehicle), should be described in simple language, avoiding technical terms and unnecessary words (with consideration of possible framing and labelling effects), with a logical sequence in presentation of material, and illustrated where appropriate with pictures, graphs or diagrams etc. Also consider interspersing the narrative with questions to involve the respondent.

Given that the issue is that of valuing health outcomes, we consider to be of utmost importance that their key dimensions should be systematically described in terms of: (i) improved health status; (ii) duration of this improvement; and (iii) probability that these improvements will occur.

Face-to-face interviews should be used (not postal or telephone surveys), with interviewers trained carefully to read material at appropriate pace, use conversational inflection, pauses and eye contact.

Pilot testing should be conducted, with qualitative research to see how respondents viewed the commodity compared to how the researcher postulated they would.

The scenarios should always be included in the report/paper so they can be viewed by readers.

3.4 Health Care WTP Studies

We reviewed studies according to: (i) the process with which it appeared they had undertaken to construct the scenario concerning the commodity being valued; and (ii) the presentation of that information in the scenario. For the first, studies were coded according to whether the scenario appeared to have been developed according to: (i) expert opinion; (ii) the researchers own views concerning the commodity, including information gained from published literature; (iii) the use of focus groups, or a delphi panel, to develop a 'consensus' view on the attributes of the commodity being valued; and (iv) the use of a mixture of the above. The results are presented in table 1.

Table 1: Scenario Development

	Frequency	Percent	Cumulative percent
Researcher derived	7	13.0	13.0
Focus group/dephi	2	3.7	16.7
Mixture researcher & expert	1	1.9	18.5
Not available	44	81.5	100.0
TOTAL	54	100.0	

Over 80% of the studies reviewed (44/54) did not provide any information on where the information concerning the commodity being valued was derived. We feel that this is the most significant result, as the scenario description, and the information upon which it is based, are the fundamentals underlying the rest of the analysis, as it is this which determines precisely what is being valued. Of the remaining 10 studies, two had scenarios based on information derived from a focus group or delphi panel, and seven were based purely on the researchers own interpretation of the attributes of the model. One study had information based on a combination of researcher and expert opinion. Thus, out of 54 studies we can only say for sure that two (4%) used information from those who have experience of the commodity. Of those studies focused explicitly on drugs, three (of nine) were developed according to the researchers own views, including literature reviews, and six provided no information.

One reason for the large number not presenting a scenario may be that over 70% (38) of studies used current or past users of the commodity as the sample valuing it (Olsen & Smith, 1999). We can see from Table 2 that the majority of those studies where the scenario development is not available where ones using users as the sample valuing the commodity (34/54). However, there are still 10 studies who did not use users (solely) who did not provide any information on scenario development. In addition, even for those using users, as was made clear in the discussion of this issue above, there is a need to determine the attributes being valued across respondents otherwise we cannot be clear that all are valuing the 'same' commodity.

Table 2: Scenario Development by Response Type

RESPONSE TYPE	SCENARIO DEVELOPMENT				TOTAL
	Researcher	Focus grp	Mixture	N/A	
Users	3	1		34	38
Conven sample	2			4	6
General pop	1		1	5	7
User+conven	1				1
User+genpop				1	1
Conven+genpop		1			1
TOTAL	7	2	1	44	54

With respect to the second issue, the presentation of the information, we reviewed studies considering whether: (i) a scenario concerning the attributes of the commodity they were valuing had been presented in the paper (by scenario we are referring to the attributes of the commodity and other features of the market, such as payment vehicle, diagnosis, and WTP question); (ii) whether just the WTP question itself was presented, but not the rest of the scenario; (iii) whether just the diagnosis was presented but not the rest of the scenario; (iv) a combination of diagnosis and WTP was presented only; and (v) whether none of these elements of the market were presented. The results are provided in table 3.

Table 3: Scenario Presentation

	Frequency	Percent	Cumulative percent
Scenario presented	26	48.1	48.1
WTP question only	3	5.6	53.7
Diagnosis only	8	14.8	68.5
Question & Diagnosis	8	14.8	83.3
Not available	9	16.7	100.0
TOTAL	54	100.0	

One can see that nearly 50% (26/54) of studies did present all elements of the scenario in some form in the paper, which is encouraging. Of those who did not, there was a roughly equal split between those providing diagnosis only (8/54), providing the diagnosis and the WTP question (8/54) and where none of this information was available (9/54). Only three studies presented just the WTP question. However, those studies focusing explicitly on drugs fared much better with respect to this issue, as can be seen in Table 4, where seven (of nine) studies presented the scenario as used.

Table 4: Scenario Presentation by Pharmacoeconomic Relevance

PHARMACO-ECONOMIC RELEVANCE	SCENARIO PRESENTATION					TOTAL
	Scenario presented	Part of the question	Only diagnosis	2 + 3	Not available	
Explicit drugs	7		1		1	9
Implicit drugs	8		1	3	3	15
Non-drug	11	3	6	5	5	30
TOTAL	26	3	8	8	9	54

Overall, that nine studies were found where none of this information is available is some cause for concern, although over time the reporting of these elements has improved, as indicated in Table 5.

Table 5: Scenario Presentation by Publication Period

PUBLICATION PERIOD	SCENARIO PRESENTATION					TOTAL
	Scenario presented	Part of the question	Only diagnosis	Part of qst +diag.	Not available	
1985-98 (5 years)	1		1	1	2	5
1990-93 (4 years)	5	1	2		4	12
1994-96 (3 years)	8	1	5	1	2	17
1997-98 (1.5 years)	12	1		6	1	20
TOTAL	26	3	8	8	9	54

WTP studies conducted in health care contain two significant problems: (i) that the reader is unable to assess where the information concerning the attributes of the commodity being valued were derived from, and thus cannot place any confidence that important and relevant dimensions have been covered⁸; and (ii) that much of the information concerning attributes and the contingent market itself is not available from the information provided in the papers. We have no way of determining whether the information was presented to respondents in a 'clear and concise' fashion, and was made understandable to them. A caveat to this would of course be that often such data is subject to space allocated in peer reviewed journals, although the variation suggests it is possible to provide a satisfactory description of the commodity and market, and present that to the reader (O'Brien, 1995). As reported in a previous paper concerning conceptual issues (Olsen & Smith, 1999), the majority of studies failed to adequately distinguish between use, option and externality values, and between the various components of health and non-health use value. This carries over here in that one cannot be sure what attributes the respondent was valuing.

⁸ Although we might say with confidence, however, that virtually none of the studies conducted any extensive assessment using focus groups or delphi panels to canvass views on the attributes of important, and pilot testing of the scenarios was not conducted, as far as we could tell, for any study.

4 Payment Vehicle

Payment vehicle refers to the manner in which the WTP payment (how WTP is expressed) is to be made in this hypothetical market. Amongst the more common payment vehicles are direct out-of-pocket (OOP) expenditure, in the same manner as most goods and services are paid for in the market, additional taxation payment, particularly for public goods, such as transport initiatives, and private insurance premium, where insurance markets dominate, such as in some health care systems (notably the USA). Although there is some debate concerning which is most appropriate for WTP studies, it is clear that WTP values derived for particular commodities are not neutral (respondents are not indifferent) to the payment vehicle employed to elicit this hypothetical payment (see, for example, Randall et al, 1978; Rowe et al, 1980; Brookshire et al 1980, 1981; Daubert and Young, 1981; Greenly et al, 1981). An inappropriate choice of payment vehicle may therefore lead to strategic bias (for example to avoid paying additional taxes) or hypothetical bias (a totally unrealistic contingent scenario, such as purchasing high technology health care OOP).

It is therefore appropriate to see the payment vehicle as forming a part of the market context, as inextricably linked to the program being valued, thus implying that there is no 'one' value for a commodity, but a series of values all dependent upon the particular payment vehicle used (Cummings et al, 1986; Rowe et al, 1980). Thus, whilst it may be desirable for the researcher to have the program valued *per se*, independent of the payment vehicle, in order to ensure a realistic market there is a need to construct a realistic payment vehicle by which the good will be purchased/provided. This inevitably means that some respondents will find the payment mechanism unattractive, and hence provide lower WTP values than for the good if valued with a different vehicle. For example, respondents may object to the form of payment outlined out of political beliefs, sense of 'fairness' or financial situation.

If the scenario is such that the respondent is taking the survey seriously, then this might result in 'protest votes', for example when they feel they are asked to pay for things which are the responsibility of others, such as 'government' in the case of health care. That is, if they already feel they are paying for it through tax they may give a low or zero bid. If the commodity is a 'private' good then the realistic vehicle is likely to be OOP or insurance, if it is a public good then taxation is more likely appropriate, and the respondent must be made aware of their responsibility, or the fact that they will be charged, some figure. However, it is clear that a plausible scenario offers a way of paying for the program that seems 'reasonable' to respondents, even if they don't necessarily like it. Which is most suitable will therefore depend ultimately upon the context in which the questions are being asked. Mitchell and Carson (1993), for example, recommend that one which is currently used for paying for the commodity, or a similar program, be used.

Within health care WTP studies there is some controversy surrounding which payment vehicle is appropriate, with some arguing that insurance premiums should be used as health care is a good which yields 'option' value in an *ex ante* situation (Birch, 1993), whilst others have argued for different vehicles (Donaldson et al, 1995). The most appropriate vehicle will also be likely to differ across cultures and countries, with insurance premium vehicles perhaps more suitable to

the USA, taxation to the UK and an increased Medicare levy to Australia. Again, one might expect the appropriate vehicle also to differ across products, with insurance perhaps more suitable for high technology items, or expensive low probability items, but OOP payment, in the form of a co-payment, for more consumer based products, such as pharmaceuticals (Drummond et al, 1997).

4.1 Recommendations

The most appropriate payment vehicle would appear to be debatable. However, in using CBA to aid priority-setting between programs it is important that the vehicle decided upon: (i) is the most realistic to respondents (which will therefore involve empirical testing of different forms of vehicle); and (ii) is used consistently by those conducting WTP studies. In terms of pharmaceuticals, it is our opinion that OOP payment is most relevant if users are asked, whereas taxation is most relevant if the general population is asked. We would therefore suggest that the relative merits of: (i) OOP payment expressed as an additional co-payment; (ii) OOP payment expressed as a brand premium; and (iii) increased taxation to fund the product at no additional direct cost to the consumer in the market, be established.

4.2 Health Care WTP Studies

We classified studies according to the vehicle used: (i) out-of-pocket (OOP); (ii) taxation; (iii) private insurance; (iv) voluntary donation; (v) combinations of these; and (vi) not available. The results are reported in table 6.

Table 6: Payment Vehicle Used

Valid	Frequency	Percent	Cumulative percent
Out of pocket (oop)	42	77.8	77.8
Tax	4	7.4	85.2
Tax+voldon	1	1.9	87.0
Oop+tax+insur	1	1.9	88.9
Not available	47.4	7.4	96.3
Oop+tax	1	1.9	98.1
Oop+insur	1	1.9	100.0
TOTAL	54	100.0	

It is clear that the most frequent payment vehicle used was OOP (42/54 studies, or 78%). Studies using multiple vehicles are those where multiple commodities are valued, and as such may be effectively excluded from analysis, suggesting that the only other real alternative form of payment was taxation, with four studies not presenting data to enable the payment vehicle to be determined. In terms of valuing pharmaceuticals, no study used a method other than OOP, as indicated by Table 7. However, it is not clear that this is necessarily the correct approach, since

none of these studies framed OOP expenditure as a copayment increase, which might create an unrealistic market.

Table 7: Payment Vehicle by Pharmacoeconomic Relevance

PHARMACO-ECONOMIC RELEVANCE	PAYMENT VEHICLE							TOTAL
	Out of pocket (oop)	Tax	Tax + vol.don.	Opp +tax +insur	Not avail.	Oop + tax	Oop + ins	
Explicit drugs	7				1		1	9
Implicit drugs	13				2			15
Non-drugs	22	4	1	1	1	1		30
TOTAL	42	4	1	1	4	1	1	54

5 Certainty Versus Risk

It has been suggested that questions in CV studies in health care should be made in probabilistic terms (Birch, 1993). This is largely because expected utility theory is based upon an explicit calculation of utility derived from the commodity in question, multiplied by the probability of receiving that utility. It has therefore been felt (also within the conduct of cost-utility analysis (CUA)) that the valuation of benefits should incorporate some element of risk or uncertainty (within CUA this has been represented by the debate concerning Standard Gamble and Time Trade-Off elicitation procedures, and latterly between Quality-Adjusted Life Years and Healthy Year Equivalents as outcome measures). However, within wider welfare economic literature, and within the environmental and transport literature concerning WTP, there is no similar recommendation, although in many cases risk has to be invoked, particularly in estimating the value of life⁹. On the contrary there is concern with incorporating probabilities for two reasons.

First, that respondents typically have difficulty facing changes in risk, particularly if those changes in risk are of a relatively small magnitude. If they are to be included, risks must clearly be of such a size as to be understandable. If respondents have little or no understanding of the probabilities presented to them, then their subjective assessment of the risk may bear little relation to the objective risk. If this is so then the objective probabilities cannot legitimately be employed. The current lower limit of understanding appears to be around 1 in 10,000 (Miller & Guria, 1991; Jones-Lee et al, 1985; Thompson et al, 1982, 1984; Thompson, 1986; Mitchell & Carson, 1989; Viscusi & Magat, 1987). Risks must also be realistic. Respondents will try to relate the objective risk presented to them to their subjective assessment of the risk. If the objective risk is perceived to be unrealistic they may provide invalid responses (Miller & Guria, 1991; Thompson et al, 1984; Mitchell & Carson, 1989).

Second, individuals may simply be so averse to risk that they exhibit what has been termed "general aversion" (Slovic et al, 1980; Kahneman & Tversky, 1979; Starr et al, 1976). Slovic et al (1980) argue that the most common way to eliminate the general anxiety which accompanies risk and uncertainty is simply to deny it. This may cause serious bias in WTP values estimated from studies involving risk and uncertainty.

However, the incorporation or not of risk will be expected to partly determine the value of the commodity, since unless respondents are risk-neutral, part of the WTP value will be to avoid (if risk averse) the risk involved (that is, valuing those commodities with a greater probability of success more than those with a lesser probability, *ceterus paribus*). It is therefore of some concern to determine the relative appropriateness of specifying benefits with certainty or risk. There are two issues to resolve here.

First, there is an issue of *ex-ante* versus *ex-post* requirement for, the probability of needing, the commodity. That is, the arguments that probability should be a feature of the CV market are largely based on the view that we require *ex-ante* views, implying that respondents *may* require

⁹ Here, for example, portraying a life saving intervention with certainty might be expected to yield an infinite value. Thus, interventions are expressed as reducing probability of death and then mathematically manipulated to estimate the WTP if the probability reduction was equal to 1 (ie certainty).

this commodity. If this is the case, then the value we are primarily determining is that of “option value” (Olsen & Smith, 1999). This may not be the value which we wish to obtain, since it will incorporate a value of risk (unless respondents are risk-neutral, which is unlikely). It also raises the issue above of whether people will misperceive the risk and thus provide a higher or lower WTP accordingly. If we do not wish to invoke such values attached to risk aversion, then it would be appropriate to value the commodity with certainty and adjust for probability separately (thus preserving risk neutrality)¹⁰. In this case we would primarily be assessing “use values” (which are likely, *ceterus paribus*, to be higher than option values, and more reflective of the actual value to the individual of the benefits conferred by the treatment) (Olsen & Smith, 1999).

Second, the presentation of probability in the successful outcome of the treatment (irrespective of whether or not the respondent has been presented with certainty in requiring the treatment). Here we are considering the attachment of probabilities to the *outcomes* of the commodity. It is likely that this probability will be both substantially higher than the one concerned with the chance of requiring the commodity, and thus perhaps overcoming the problem of levels of understanding, and the valuation of attitude toward risk may well in such a circumstance be of relevance to the valuation of health care programs.

5.1 Recommendations

Recommendations are difficult to find until there is some general indication of whether the value of risk aversion is seen as a ‘legitimate’ source of benefit resulting from health care programs. However, it is our opinion that probability should be included at the level of program success in achieving the proposed outcome/benefits. It is not our opinion that all programs should be presented on an *ex-ante* basis, unless option value is the desired valuation, as it will not capture the full value of the program as gained by those who use it. In addition, it is suggested that respondent understanding of probability is limited, and so incorporating as little as possible will reflect this. Furthermore we would recommend that whenever probabilities of any type are incorporated the valuation is undertaken under varying probability levels to assess the degree of understanding of probability (that a larger probability should yield a higher value) and an estimation of the degree of risk aversion.

5.2 Health Care WTP Studies

We found that 38 studies (70%) did not have any element of risk or uncertainty presented in the contingent market (including six (of nine) explicit drug studies). This is a mixture of *ex-ante* versus *ex-post* risk, and risk in the benefits to be derived from the commodity - this split is approximately 50:50, with around eight studies presenting the benefits of the commodity with probability attached. Table 8 presents a cross-tabulation of the incorporation of risk tabulated with the type of respondent used. One can see that nine studies used users and presented risk within the scenario. It is arguable whether a user will respond to the presentation of risk, and not simply place certainty on the equation given their status viz. the commodity in question.

¹⁰ In comparison, the approach generally adopted in estimating QALYS is to present the likelihood of requiring the intervention as certainty, and attach probabilities during subsequent analysis.

Table 8: Risk/Problems in the Scenario by Respective Type

RESPECTIVE TYPE	RISK/PROBLEM IN THE SCENARIO		TOTAL
	Yes	No	
Users	9	29	38
Conven sample	4	2	6
General pop	1	6	7
User+conven	1		1
User+genpop	1		1
Conven+genpop		1	1
TOTAL	16	38	54

6 Time Period for the WTP Estimate

There is little discussion within the literature (environmental, transport or health care) relating to the time period over which WTP values could, or should, be assessed. Typically within environmental and transport economics the values are for a one-off payment over a life-time for a specific intervention. Occasionally, the payment will be yearly for a set number of years, or indefinitely. However, what can we say about the appropriateness, or comparability, of differing time periods?

WTP measures, as typically used in environmental economics and transport economics, and to value life, are sometimes referred to as 'overall' or 'lifetime' measures. That is, the individual is asked for their maximum WTP for a 'rest-of-life' shift in the state of the world, such as a shift in health status or health risk (Johansson, 1995). In contrast, WTP studies of health care programs typically calculate WTP according to a specified time period, usually somewhere less than or equal to one year. However, there is no *a priori* reason to suspect that these two measures (lifetime and present value of incremental time period WTP expressions) will yield equivalent WTP values, or indeed even the same sign (Blackorby et al, 1984).

The 'life-time' model assumes that the individual will be free to borrow (or lend) any amount of money required at the prevailing market rates. In contrast the 'annual' model assumes that the individual will be constrained by his or her 'annual' income, and will thus be unable to adjust consumption expenditure and borrowing over time. Under these conditions it can be shown that the (discounted) value of the sum of 'annual' WTP values will be less than the maximum 'life-time' WTP value (Johansson, 1995). In general it can also be shown that if the sum of the discounted 'annual' WTP values is positive then the life-time WTP value will also be positive (although by a larger actual dollar value). However, this is ambiguous in certain circumstances where the present value of 'annual' WTP values may be negative, yet 'life-time' WTP is positive (Blackorby, 1984).

Any conclusion concerning time period to be used would seem to depend upon two factors. First, whether the commodity being valued is *amenable* to payment on, for example, a yearly basis or can only be assessed as a one-off payment. Second, whether the *policy question* relates to such life-time one-off payments or to annual payments. For example, it is clear that for a commodity to be funded by taxation (used as the payment vehicle), annual payment periods are more appropriate than one-off payments to the individual. We might conclude therefore that the appropriate time period is to be that which corresponds to the actual payment system invoked in the contingent market. It is possible, however, that this will make comparison difficult, for the reasons mentioned, as in comparing two commodities identical in all ways (including the 'true' value of the benefits), with one valued on a once-off basis and one on an annual basis, the expressed WTP would differ solely due to the time period for payment. An additional factor is that respondents, in general, do not appear to feel any commitment to a multi-year payment beyond a couple of years (Mitchell & Carson, 1993). As such, if annual payment is to be invoked then expression as 'multi-year' payment is to be avoided, and perhaps only used in cases where it is clear that the provision of the program would cease if regular payments were not continued.

6.1 Recommendations

On the basis of this evidence we would conclude that there needs to be some consistency within the time period used in WTP studies to enable comparability between them. We would therefore suggest that one of two ways of framing the question, with respect to time period, be adopted: either as (i) a one-off payment but constrained to within any one year, by invoking the respondents annual (rather than lifetime) income, or as (ii) a regular yearly payment, with the value derived for ‘this year’ only, not for a ‘hypothetical’ year.

6.2 Health Care WTP Studies

We classified studies according to time periods of: (i) 12 months; (ii) 3 months; (iii) 1 month; (iv) 1 week; (v) lifetime (explicit); and (vi) per intervention (no time period, implying a one-off payment each time the commodity is used). The results are presented in Table 9.

Table 9: Time Period for WTP Question

	Frequency	Percent	Cumulative percent
12 months	10	18.5	18.5
3 months	1	1.9	20.4
1 month	14	25.9	46.3
1 week	1	1.9	48.1
Lifetime	2	3.7	51.9
Per intervention	26	48.1	100.0
TOTAL	54	100.0	

The majority of studies expressed the WTP question in terms of per intervention, as one-off payments each time the commodity is consumed. It is unclear if a time period is invoked here, as respondents may be viewing expenditure in terms of annual income, or may be considering what they could borrow over a life time in order to consume the product. Out of the six studies invoking a budget constraint, three of those used a per intervention payment, two a per month payment and one a yearly payment. Thus only three of 26 studies seeking an open ‘per intervention’ payment reminded respondents of budget constraints, and therefore would imply a time horizon.

Of those invoking a specific time period, the most frequent used was the one month time period, followed by the single year period. Few studies used any other period. Studies focused explicitly on drugs, however, demonstrated a different pattern, as illustrated in Table 10. Of nine such studies, the majority (five) used a time period of one month.

Table 10: Time WTP by Pharmacoeconomic Relevance

PHARMACO-ECONOMIC RELEVANCE	TIME WTPQ						TOTAL
	12 mth	3 mths	1 mth	1 week	Lifetime	Per int.	
Explicit drugs	1	1	5		1	1	9
Implicit drugs	1		8			6	15
Non-drug	8		1	1	1	19	30
TOTAL	10	1	14	1	2	26	54

Another aspect of interest is that the profile of time periods used in the analysis of WTP differs slightly to that asked of individuals. The areas of difference are two fewer studies analysed over a one month period than were assessed over that period, and one per intervention less, with correspondingly one more over a lifetime and two analysed over both a yearly and monthly period. That is, a few studies are *transforming* data from the time period which was used by respondents to assess their WTP, and the period used in the analysis of the value of that commodity (although for the nine explicit drug studies none used a different time period in analysis than asked of respondents). We return to this issue of transformation later.

7 Hypothetical Bias

“I see that saying and doing are two different things, and hereafter I shall better observe this distinction”

John Bunyan, *The Pilgrim’s Progress*.

“Ask a hypothetical question and you get a hypothetical answer”

Scott (1965, p37).

In recent years there has been some emerging evidence that WTP based on hypothetical CV studies is higher than WTP revealed in actual decisions. In this case individuals are systematically overestimating their WTP in such CV studies, which, it is argued, may be due to their ‘hypothetical nature’ (Johannesson, 1996)¹¹.

The CV method has been criticised for eliciting hypothetical answers as a necessary result of the hypothetical survey situation (Diamond & Hausman, 1993). It is argued that “a single purchase of an unfamiliar commodity represents a guess as to what the commodity might be worth, rather than an evaluation based on experience” (p13, Diamond & Hausman, 1993). The argument follows that CV surveys are conducted because products being valued are not available in the market. As such they are either unfamiliar, or are familiar but presented in unusual, and unfamiliar, circumstances. In this case individuals will have no experience of the product or close substitutes and will therefore have no underlying information or preference upon which to base their decision concerning its value to them.

Evidence to support this view is gained from considering the information attained by marketing firms and departments concerning predicted markets prior to new product launches. It is suggested that about 65% of all new product launches fail precisely because pre-launch marketing suggested a significant market at the proposed price, with the company only to find that after launch this market was non-existent¹². Another supportive body of evidence concerns the use of opinion polls prior to voting in elections or other referenda. Diamond & Hausman (1993) argue that such polls are frequently inaccurate, where the pre-voting polls can often provide an indication of the result being in favour of one proposal, only to find that the actual vote brings in the opposite result.

For economists, amongst other professions, there is a distinction between attitudes and behaviour. For example, to observe that someone likes a Mars Bar is not the same as them actually purchasing and eating it. Even for them to state that they intend to eat it is not the same as actually doing so. It is only when the individual actually eats it that the link between intent and behaviour can be made. Behaviour is ultimately what is observed, and it is argued that it is a long road from intent to acting out that intent (Bishop & Heberlein, 1986).

¹¹ Although it is debatable whether hypotheticality is a separate and distinct bias, or just encompasses other biases, such as strategic bias, which occur **because** of the hypothetical nature of the exercise.

¹² There is a considerable amount of evidence relating to these points, concerning consumer purchasing anticipations, new product launches and expressed demand for proposed new transport services. This body of evidence is neatly summarised by Kemp and Maxwell (1993).

Another version of hypothetical bias has been termed ‘noncommitment bias’. This concept simply points out that purchases in the market are characterised by ‘removing real dollars from one’s wallet’, whereas responses to a survey do not. Thus, survey responses do not require the same level of financial commitment as a ‘real’ purchase, and so the responses may be easily swayed by a variety of factors, such as wishing to get the interview over with quickly or a wish to appear reasonable (Kemp & Maxwell, 1993).

However, according to Mitchell and Carson (1993), what matters is the difference between ‘formal hypotheticality’ and ‘substantive hypotheticality’. The former are forms of experiment, commonly thought of as containing phrases such as ‘imagine’, ‘what if’, ‘suppose’ or ‘presume’. This framing is therefore suggestive to respondents that the choice they are offered is of little *real* consequence, mitigating against obtaining an economically meaningful response to the WTP question. However, the latter is a situation where, although the respondent knows that they are not *actually* at that point being asked to pay, the situation is *realistic* enough for them to believe this *could* happen. This involves simulating a plausible real-life situation with sufficient possibility that respondents take it seriously. An example might be the form of payment vehicle. Respondents are not used to purchasing most health care OOP, but through taxation, leading to a conclusion that perhaps taxation is the most realistic means of payment vehicle¹³.

We suspect that the conclusion to this issue is that it is true that a hypothetical question will yield a hypothetical answer, but that does not mean it is not of any use in public decision making. If a well-constructed question is asked, people will try to oblige with honest answers to the best of their ability. Whilst these may not *exactly* match actual behaviour in a market, they are the best indication of the individual intent and value of the program, and in the absence of the market this is perhaps the best we can hope to achieve. Following from this, practically this means that the market context within which the commodity is offered and valued should be as neutral as possible, in the sense that it should be realistic but not likely to elicit strong preferences due to the contextual features itself. It has considerable overlap with the choice of payment vehicle (for example, in the UK CV surveys using private market provision with insurance based payment vehicles may be seen as unrealistic and/or provoke strong anti-market or pro-NHS feelings from respondents which will be registered in their WTP, such as in protest votes). Lack of realism in the market context will be at risk of “hypothetical bias”, where respondents will give unrealistic answers to unrealistic situations.

¹³ The fact that people, in general, resent paying taxes might also mitigate against excessively high WTP values ensuring that the situation is taken seriously and that results will err on the side of caution.

7.1 Recommendations

It is clear from the above discussion that the 'hypothetical' market must be as real as possible. The CV market should therefore be presented in a realistic manner with respect to the policy change being valued, payment vehicle and likely benefits. Respondents, we believe, want to give realistic values, and as researchers we should try to enable them to do so by presenting clear and considered information in a manner which appears realistic (ie appears that it could actually happen in practice).

With respect to pharmaceutical evaluation this would seem to be a great deal easier to achieve than, for example, valuing environmental amenities. The products, or at least similar products, have been bought, and therefore valued, before. Furthermore, there is a 'market' in existence with which respondents will have had experience. It therefore seems sensible that the contingent market should resemble this as closely as possible (which of course impacts upon payment vehicle, and time period for payments etc). One note of caution, however, would be that such familiarity has likely produced a 'value' which many will place on products reflecting the 'price' to them in the existing market for similar products. This may act as a form of 'first-response' bias, or may yield strategic bias - both of which are considered in more detail below.

7.2 Health Care WTP Studies

Not applicable.

8 Strategic Bias

Strategic bias is said to exist where a respondent *deliberately* gives a WTP amount that differs from their true WTP amount (conditional on the perceived information) in an attempt to influence the provision of the good and/or the respondents level of payment for the good, and is believed to be a potentially significant problem in estimating 'true' WTP values (Mitchell & Carson, 1989).

The strategic bias problem can be easily summarised as making the exercise *sufficiently* realistic to elicit truthful values, but not *too* realistic to elicit 'free-riding' behaviour. This is not an insubstantial undertaking. If the respondent feels that they will indeed be charged and expected to pay in practice their proffered WTP amount, yet feel that the provision of the good is largely independent of their valuation (either that others will provide a sufficiently high WTP amount to ensure provision, or that this is in effect a study to determine pricing of a commodity which has already been decided to proceed with) then the incentive is to underrepresent their true WTP and 'free ride' by being ensured provision of the good, yet being charged an amount less than they would actually be WTP (they therefore receive a substantial 'consumer surplus'). Alternatively, if the respondent believes that the WTP value given is to determine whether the commodity is provided, yet they will most definitely not have to pay that amount in practice, the incentive is to give an inflated WTP, higher than their true valuation, in order to ensure the service is provided, confident that they will not have to pay the amount specified.

Empirical evidence is equivocal concerning the presence of strategic behaviour, and largely depends on how well the CV market is specified to respondents, although the weight of evidence suggests that its effect, if present, is negligible (Groves, 1973; Clarke, 1971; Groves & Ledyard, 1977; Tidman & Turlock, 1976; Rowe et al, 1980; Bohm, 1972; Scherr & Babb, 1975; Brookshire et al, 1976; Smith, 1979; Milon, 1989; Jones-Lee, 1989). It has also been argued that discrete questionnaires (see Smith, 1997) reduce the incentive for strategic bias (Mitchell and Carson, 1989; Carson 1991). However, it appears that it is not the *incentives* for strategic bias which differ, but the *potential* for such strategic behaviour to affect the results (Johannesson, 1996). This is because, in such a binary question, the respondent only gives a yes or no response to an offered bid, thus having little effect on the overall average bid level. However, if *all* respondent behaved in a strategic fashion then the results of the survey would be affected in the same manner as an open questionnaire.

However, it is rational, and to be expected, that respondents will provide an overestimate of their WTP for a good such as a pharmaceutical. This is because respondents know the system by which the pharmaceutical will be introduced, in terms of the PBS and co-payment system currently in operation. Thus there is a two-stage decision making process involved. The first stage will be the WTP valuation survey, where the respondent is asked for the valuation of the pharmaceutical, in full knowledge that given the system in operation they will never have to pay that price. Their incentive in this situation is to overestimate their WTP to ensure the pharmaceutical becomes available. The respondent will be expected to behave in the way to ensure that the drug becomes available as an option to them should they ever need to use it. The second stage is the actual decision facing the consumer in practice which is to buy or not the

pharmaceutical once it has come to market at the market price, which one expects to be lower than the WTP amount they suggested - especially given a system like the PBS which ensure a low maximum fee for the respondent to pay in practice. It is therefore important that if WTP studies are to be undertaken for pharmaceuticals that this is recognised and the study adjusted to try to account for this bias. For example, by expressing WTP as an addition to the user fee, or a 'brand premium'.

An issue within recent health care studies, however, has been the tendency, for UK authors in particular, to explicitly state to respondents that they will definitely not have to pay for the commodity being valued. According to the evidence above it is likely that this sort of statement will reinforce some element of hypotheticality or strategic behaviour - if one wishes to see the commodity funded by the health service, but are told one is definitely 100% guaranteed not to have to pay for it, then the incentive is to give as high a value as is realistically possible to skew the results towards favouring the commodity.

8.1 Recommendations

Although there is evidence to suggest that WTP responses *may* be made in a mendacious or random manner, such that subjects may seek deliberately to misrepresent their true preferences, the weight of evidence suggests respondents do not *seek* to do this. We would, however, conclude that it is important as a *potential* issue, and would recommend that respondents are not told they will not have to pay for this product, but rather the contrary and informed that this hypothetical situation is one which has a substantial probability of occurring. Any incentive would, in this situation, be toward free-riding, which we believe is preferable to an incentive to over-estimation, as it provides a conservative estimate of the value of the commodity (as recommended, for example, by the NOAA Panel (1994)).

8.2 Health Care WTP Studies

Strategic bias was tested for in 11 studies (20%), and in each case was not found to be a significant bias. However, the majority (80%) did not test for the existence of such bias and so we cannot be confident that it was not present in the majority of studies conducted. Of the 11 studies testing for bias, three were explicit evaluations of pharmaceuticals, and one implicit.

9 Conclusions and Recommendations

This review provides a comprehensive overview of the major issues involved in the construction and specification of the contingent market for WTP studies. 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.

In terms of constructing the contingent market, we believe this to be the key methodological area, as it forms the foundation for the resultant WTP values. Although it is difficult to provide rigid recommendations for how such studies should be done, several key features do present themselves from our reading of the literature.

- 1) There is a requirement to have appropriate discussion with those commissioning the study; why is an evaluation of this sort felt desirable and what is the precise nature of the program to be valued. This is vital to ensure that the economic issues and nature of the program are clarified.
- 2) Focus groups of general public and/or specifically involved sub groups should be held. These wide-ranging discussions explore how respondents view the program, its context and attributes. This is an iterative process as refinements to the scenario and market context can be fed back to such groups for refinement and clarification.
- 3) The initial draft scenario and instrument should be piloted with in-depth interviews on aspects of the respondents thoughts and comments on undergoing the survey.
- 4) Once refined the survey should then be conducted on small groups by professional interviewers. This test the instrument 'in the field', and allows further refinement subject to any difficulties faced by the interviewers.
- 5) Once the researchers are happy with the instrument, the pilot tests should be conducted on samples of reasonable size to allow quantitative data to be collected to allow identification and judgement concerning covariates which may predict WTP responses - for discrete surveys this stage is to be used to identify the appropriate bid vector. This pilot stage may be iterative depending upon the initial quality of the instrument.
- 6) Finally, the survey goes 'live' with the researchers confident of obtaining valid and reliable WTP values.

In summary, these recommendations concerning the market context may be simply expressed in terms of five specifics required of any CV study:

- 1) Specific context: eg. "Keeping in mind your actual number of episodes of migraine"
- 2) Specific action: eg. "what is the most you would personally be WTP"
- 3) Specific vehicle: eg. "as a direct out-of-pocket copayment"
- 4) Specific time frame: eg. "each year"
- 5) Specific target: eg. "to reduce the time your migraine lasted from 2 days to 1 day".

Reviewed against these recommendations, it is clear that WTP studies conducted to date have, overwhelmingly, performed poorly in the construction, specification and presentation of the contingent market. 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 construction of a contingent market to be debated and agreed upon if such studies are going to become a useful addition to the economic evaluation 'toolkit'.

REFERENCES

Acton JP 1973, Evaluating public programs to save lives: The case of heart attacks, Santa Monica, RAND Report, R-950-RC.

Ajzen I, Fishbein M 1977, 'Attitude-behaviors relation: A theoretical analysis and review of empirical research', *Psychological Bulletin*, vol 84, pp 888-918.

Appel L.J, Steinberg EP, Powe NR, AndersonGF, 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 an 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.

Birch S, Gafni A. 1993, 'Changing the problem to fit the solution: Johannesson and Weinstens (mis)applciation of economics to real world problems', *Journal of Health Economics*, vol 12; pp 469-76. See also: Gafni A, Birch S 1993, 'Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem', *Canadian Medical Association Journal*, vol 148, no 6; pp 913-917.

Bishop RC, Heberlein TA 1986, 'Does contingent valaution work?', in *Valuing Environmental Goods. An Assessment of the Contingent Valuation Method*, eds RG Cummings, DS Brookshire, Schulze WD, Rowman & Allanfield, New Jersey.

Blackorby C, Donaldson D, Moloney D 1984, 'Consumer's surplus and welfare change in a simple dynamic model', *Review of Economic Studies*, vol 51, pp 171-176.

Bohm P 1972, 'Estimating demand for public goods: An experiment', *European Economic Review*, vol 3, pp 111-130.

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.

Carson RT 1991, 'Constructed markets', in *Measuring the Demand for Environmental Quality*, eds JB Braden & CD Kolstad, Amsterdam, Elsevier/North Holland.

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.

Clarke EH 1971, 'Multi-part pricing of public goods', *Public Choice*, vol 11, pp 17-33.

Cummings RG, Brookshire DS, Schulze WD 1986, *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Rowman & Allanheld, New Jersey, USA..

Daubert JT, Young RA 1981, 'Recreational demands for maintaining instream flows: A contingent valuation approach', *American Journal of Agricultural Economics*, vol 63, pp 666-676.

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, Thomas R, Torgerson DJ 1995, Open -ended versus payment scale approaches to eliciting willingness to pay, Discussion paper 01/95, Health Economics Research Unit, Departments of Public Health and Economics, University of Aberdeen, Aberdeen, Scotland.

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, Johansson 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.

Fischhoff B, Furby L. 1988, 'Measuring values: A conceptual framework for interpreting transactions with special reference to contingent valuation of visibility', *Journal of Risk and Uncertainty*, vol 1, pp 147-184.

Gerard K., Dobson M, Hall J 1992, 'Framing and labelling effects in health descriptions: Quality of life years for treatment of breast cancer', *Journal of Clinical Epidemiology*, (submitted).

Greenley DA, Walsh AG, Young RA 1981, 'Option value: Empirical evidence from a case study of recreation and water quality', *Quarterly Journal of Economics*, vol 95, pp 657-673.

Groves T 1973, 'The incentives in teams', *Econometrica*, vol 41.

Groves T, Ledyard J 1977, 'Optimal allocation of public goods: A solution to the 'free rider' problem', *Econometrica*, vol 45, no 4, pp 763-809.

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

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, 'The contingent valuation controversy in environmental economics and its relevance to health services research', *Journal of Health Services Research and Policy*, vol 1, pp 116-117.

Johannesson M, Johansson P.O, 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.

Kahneman D, Tversky A. 1979, 'Prospect theory: An analysis of decisions under risk', *Econometrica*, vol 47, no 2, pp 263-291.

Kahneman D, Ritov I, Jacowitz KE, Grant P 1993, 'Stated willingness to pay for public goods: A psychological perspective', *Psychological Science*, vol 4, pp 310-315.

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

McGuire A, Henderson J, Mooney G 1988, *The Economics of Health Care*, Routledge; London.

McNeil BS, Panker J, Cox HC, Tversky A 1982, 'Patient preferences for alternative therapies', vol 306, pp 1259-1262.

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.

Mitchell RC, Carson RT 1988, 'How far out on the learning curve are we with contingent valuation?' in *Proceedings of the Conference on the Role of Social/Behavioral Science in Water Resource Management*. eds AD Baumann, YH Haimen, New York, American Society of Civil Engineers.

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

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.

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

Randall A, Grunewald O, Pagoulatos A, Ausness R, Johnson S 1978, 'Reclaiming coal surface mines in central Appalachia: A case study of the benefits and costs' *Land Economics*, vol 54, pp 472-489.

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.

Scherr BA, Babb EM 1975, 'Pricing public goods: An experiment with two proposed pricing systems', *Public Choice*, vol 23, pp 35-48.

Schoemaker PJH 1982, 'The expected utility model: Its variants, purposes, evidence and limitations', *Journal of Economic Literature*, vol 20, pp 529-563.

Scott A 1965, The valuation of game resources: Some theoretical aspects, Canadian Fisheries Reports, no 4.

Slovic P, Fischhoff B, Lichtenstein S 1980, 'Facts and fears: Understanding perceived risk', in *Societal Risk Assessment: How Safe is Safe Enough?*, eds RC Schwing, WA Albers, New York, Plenum Press, pp 181-216.

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

Smith RD, Dobson M 1993, 'Measuring utility values for QALYs: two methodological issues', *Health Economics*, vol 2, no 4, pp 349-355.

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*.

Pauly MV 1995, 'Valuing health care benefits in money terms', in *Valuing Health Care*, ed FA Sloan, Cambridge University Press, Cambridge.

Smith RD, Olsen JA, Harris A 1999, A Review of Methodological Issues in the Conduct of WTP Studies in Health Care II: Administration of a CV survey. *CHPE Working Paper No. 85*.

Smith RD, Olsen JA, Harris A 1999a, 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*.

Smith VL 1979, *Research in Experimental Economics*. Vol 1, Connecticut, JAI Press.

Starr C, Rudman R, Whipple C 1976, 'Philosophical basis for risk analysis', *Annual Review of Energy*, vol 1, pp 629-662.

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.

Thompson MS, et al 1982, 'Willingness to pay concepts for societal decisions in health', in *Values and Long Term Care*, eds RL Kane & RA Kane, Lexington (MA), Lexington Books .

Tidman T, Turlock G 1976, 'A new and superior process for making social choices', *Journal of Political Economy*, vol 84.

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

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

Viscusi WK & Magat WA 1987 *Learning About Risk*, Cambridge (MA), Harvard University Press.

APPENDIX 1

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 '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 1996, 'Measuring heart patients' willingness to pay for changes in angina symptoms', *Medical Decision Making*, 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, 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 & 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, 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, 1994; 3: 509-513.

Neumann PJ, Johannesson M. 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 & 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 paediatric facial lacerations', *Journal of Paediatrics*, 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. 'Willingness to pay for poison control centers', *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 & 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, 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, et al 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 M 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 & Medicine*, vol 43, pp 985-998.

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