Who Have Been Asked to Value What?

A Review of 54 ‘Willingness-to-Pay’ Surveys in Healthcare

Professor Jan Abel Olsen
Visiting Professor, University of Tromso, Norway

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

April, 1999
ISSN 1325 0663
ISBN 1 875677 95 X
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 Public Health and Community Medicine at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

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 receives core funding from the National Health and Medical Research Council and 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 is a result 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.

The paper was originally presented at the Summer 1998 meeting of the United Kingdom Health Economics Study Group held in Galway, Ireland. A thorough revision of this paper will be presented at the forthcoming Second International Health Economics Association Conference, to be held in Rotterdam in June 1999.
The aims of this paper is to outline three types of arguments put forward that WTP is superior to QALYs, and to review how empirical studies adhere to their implications. The first, that WTP is the ‘theoretically correct’ approach, because of its foundation in welfare economics, is being dismissed, as it is no argument. The second, that WTP imposes no restrictions as to which attributes of a programme people are allowed to value, makes sense. The paper will therefore focus on an inquiry into the scenario descriptions in the surveys. The third argument is the cost-benefit view that WTP can assist in improving social efficiency. We argue that it is impossible to infer from a partial WTP study (38 studies were partial) that a new programme should be funded from a given total health care budget. This we cannot tell without knowing the opportunity costs in terms of the benefits forgone from the displaced programmes.
1 Introduction

2 ‘WTP is Theoretically Correct’
   2.1 ‘To Each According to Need’ or ‘To Each According to Utility’

3 Practical Applications of WTP in Health (Care)
   3.1 Method for Selecting Papers to Review
   3.2 Exclusion Criteria
   3.3 Review Process
   3.4 Characteristics of the Papers

4 ‘WTP Enables More Comprehensive Valuations of Benefits’
   4.1 The QALY - A Restrictive Measure of ‘Utility’
   4.2 Beyond Health Outcomes
   4.3 WTP: Use Value, Option Value and Externalities
   4.4 How Have They Been Asked?
   4.5 Valuing What?

5 ‘WTP can Advise on Overall Social Efficiency’
   5.1 WTP in Practice – Partiality, Implications And Mode Of Payment

6 Two Health Policy Issues: Equity and Transferability of Results
   6.1 Equity
   6.2 Limited Transferability of Outcome Values

7 Conclusion

References
Appendix 1: The Complete Reference List of Papers Reviewed

Appendix 2: Classifying into Main Types of Journals

List of Tables

Table 1: Publication Period by Type of Journal
PUBPER * Type of Journal Cross-tabulation

Table 2: Different Values by Type of Respondents
RESPTYPE * Different types of Value Cross-tabulation

Table 3: Data Collection by Type of Respondents
RESPTYPE * DATACOLL Cross-tabulation

Table 4: Degrees of Comprehensive Outcome Descriptions

Table 5: Degrees of Comprehensive Description of Non-Outcome Attributes

Table 6: Main diagnoses by Pharmacoeconomic Relevance
DIAGNOSI * PHARMREL Cross-tabulation

Table 7: Partiality by publication period
PUBPER * PARTIALI Cross-tabulation

Table 8: Implication by Type of Journal
IMPLICAT * Type of Journal Cross-tabulation

Table 9: Implication by CBA-Measure
IMPLICAT * CBAMEASU Cross-tabulation

Table 10: Mode of Payment

Table 11: Recognition of (inequitable) distribution by publication period
PUBPER * Recognition of Distribution Cross-tabulation
A Review of 54 ‘Willingness-to-Pay’ Surveys in Healthcare

1 Introduction

The recent literature on economic evaluation of health care has shown increasing interest in the use of willingness-to-pay (WTP) as a measure of health benefits. There appears to be three types of arguments put forward that WTP is superior to quality-adjusted-life-years (QALYs) in the assessment of health outcomes. First, WTP is the ‘theoretically correct’ approach, because of its foundation in welfare economics. Second, WTP imposes no restrictions as to which attributes of a programme people are allowed to value, as opposed to QALYs where only health states are being valued. Third, benefits are valued in the same unit as costs which is required for advising decision makers on improvements in allocative efficiency.

The aim of this paper is to outline each of these arguments, and review how empirical studies adhere to their implications. The following section discusses the assertion that ‘WTP is theoretically correct’. Section 3 describes the method for selecting the 54 papers reviewed and gives some characteristics of the surveys. Section 4 deals with the argument that WTP enables more comprehensive valuations of benefits than QALYs, and provides a careful review of the scenario descriptions in the surveys. The major focus of this review lies on which types of benefits are being valued, and on the comprehensiveness of the scenario descriptions. As such this review has a different focus from the recent review by Diener et al (1997). Section 5 discusses ‘the cost-benefit argument’ that WTP can assist on improving the allocative efficiency. When reviewing the partial nature of the surveys conducted, it is questioned whether the majority of the studies could provide much help on this matter. Prior to concluding, Section 6 addresses two issues which - we believe - would be of particular importance for public health policy.

We would like to emphasize very strongly that the current paper is indeed ‘work in progress’. We therefore take the privilege of using this fact as our explanation for what we may have misinterpreted in the reviewed papers.

2 ‘WTP is Theoretically Correct’
Contingent valuation (CV) involves asking individuals directly in a hypothetical survey the maximum amount they are willing-to-pay (WTP) to have the commodity in question, or the minimum amount they would be willing-to-accept (WTA) in compensation to be deprived of it. Among the various theoretical measures which exists (see O’Brien and Gafni 1996 for an overview of these measures), WTP is the most widely applied. As all the 54 CV-studies in health reviewed here use WTP, except for one which used WTA, they will henceforth be signified simply by the term WTP.

Virtually all economists would accept that WTP has its foundation in neoclassical welfare economics. However, it is one thing to acknowledge that WTP has this theoretical basis, and in that sense is 'theoretically correct', but it is another to use that as an argument per se for applying it to health and health care. A technique either does, or does not, accord with the principles outlined within a theory. In this sense, as judged from whether being based on neoclassical welfare economics, WTP is indeed superior to QALYs. The argument for the superiority of WTP in this context has been made more on the ‘normative’ grounds that it is right because it holds to the right theoretical foundation, i.e. neoclassical welfare economics. Pauly (1995) appears to go even further in holding that WTP represents the only benefit measure with a basis in economic theory. The question is whether this is the theoretical basis we wish to use in the allocation of health care resources.

To hold that something is right because it accords with a theory is more religious than scientific. However, it is surprising to experience - in texts as well as in discussions - how frequently many economists justify the appropriateness of WTP as a model by reference to its link to welfare economics. The rightness of an evaluation approach is not to be judged from its disciplinary basis (economics), nor from its theoretical foundation (neoclassical welfare economics). Rather, it is to be judged on the basis that its premises and value judgements correspond with wide commonsense; a method is ‘correct’ if society wants to use it as a basis for allocating health care.

While sometimes acknowledged, most economists would have been taught that this theory of welfare economics is normative. The crucial issue then is the extent to which the value judgements correspond with those of the institutions whose decision makers WTP surveys are supposed to aid. In a private insurance market or in patient payment markets where health care is distributed according to how individual preferences are expressed through WTP, there seems to be no clash of values. In cases of market failures, WTP would be used as a way of ‘constructing a market’.

Interestingly, in the areas where WTP studies were initially applied, the method sought to construct markets for goods which had never been offered in a market. Health care is different in that markets have been deliberately destructed in many countries for equity and efficiency reasons. The background for public health services is to distribute health care according to need. Is then WTP a potentially appropriate measure of need? If not, there is a clash of values between welfare economics and public health policy.

2.1 ‘To Each According to Need’ or ‘To Each According to Utility’

The inferral of individual values to public policies have two dimensions. First, it is assumed that all preferences which are signaled through WTP are relevant in the given social context. No inquiry is made into the social importance of the source of the utility, i.e. the distinction between needs and desires. If a stated WTP for a particular intervention reflects an
individual's desire for pleasant amenities and appealing process of care, that is to be dealt with as an equally legitimate source of utility as if the same WTP were to reflect the value of a health improvement. Of course, if the health service aims to respond to individual preferences, there is no clash of values. However, if the health service aims to meet people's 'health needs', WTP may not be an appropriate way of signaling individual variations in intensities in needs, because not all preferences are relevant in terms of being related to health needs. It is a fundamental philosophical issue to question whether all individual preferences should be counted unfiltered in a social context. Which preferences are there for the health service to meet? Should the government fund in proportion with utility, or in proportion with need? While these questions clash with the concept of consumer sovereignty, it is still the case that most publicly funded health services are rooted in needs rather than preferences.

Secondly, it is assumed that all relevant preferences are signaled through one's WTP. No distinction is made between the individual as consumer and the individual as citizen. The WTP method asks for subjective preferences ('how much are you WTP?') and not ethical preferences related to the person's view on what should be the basis for priority setting in the public health service. The individual is implicitly assumed to have no preferences of relevance for this decision making beyond those which he signals through his WTP. There is evidence that the priority setting implied from individuals’ WTP figures differ from what is expressed when doing explicit ranking of the same programmes (see eg Olsen 1997).
3 Practical Applications of WTP in Health (Care)

3.1 Method for Selecting 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.

3.2 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.

3.3 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:

- Year of publication
- Journal type of publication (economics, health economics, policy or clinical)
- Respondents (users, convenient sample, general population)
- Data collection method (face-to-face, telephone, postal, self-administered)
- Non-use value assessed (option value or externalities)
- Data collection method (face-to-face, telephone, postal, self-administered)
- Outcomes assessed (health, duration, probability, and an assessment of the comprehensiveness of the outcomes valued)
- Non-outcomes assessed (care, information or process utility aspects)
- Presentation of the scenario description (separate or as part of the question)
- Recognition of risk or probability in the scenario
- Primary diagnosis
- Pharmacoeconomic relevance (explicit, implicit or no drug)
- Partiality of analysis (one intervention, close substitutes, different diagnoses)
- Methodological and/or policy implications considered
- Payment vehicle (out-of-pocket, tax, insurance)
- Measure of CBA (NPV or just WTP value)
- Equity considerations

The above criteria are used as the empirical basis within this paper. (A review of the methodological aspects will be dealt with in a forthcoming paper). Both authors independently classified each study along these criteria, and then conferred to assess the degree of agreement. Areas upon which the authors had differed were then reviewed and a consensus decision made.
3.4 Characteristics of the Papers

There appears to be a high concentration of authorship, dominated by four authors who have been involved in 29 surveys (two of them in 23 studies, Donaldson and Johannesson). Partly as a consequence of this, there is also a concentration in terms of the country of origin: 19 from the US, 13 from Sweden, 12 from the UK, and 4 from Canada. Hence, three countries have more than 80% of the studies.

There is a large increase in the number of published papers, from 1 per year in the period 1985-1989 to 12 in 1997 (and 8 published or forthcoming that we have been able to identify within the first five months of 1998). While there is a huge relative increase, the absolute number of WTP studies is still minute as compared with the 1456 CEAs and 450 CUAs published until 1997 (OHE-HEED database).

As reported in Table 1, the growth in the number of published studies have been in the clinical and in the public health journals.

Table 1: Publication Period by Type of Journal

<table>
<thead>
<tr>
<th>PUBPER 1985-89 (5 years)</th>
<th>Clinical</th>
<th>Public health</th>
<th>Health Econ</th>
<th>Econ</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985-89 (5 years)</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>1999-93 (4 years)</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>1994-96 (3 years)</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>1997-98 (1.5 years)</td>
<td>4</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>TOTAL</td>
<td>14</td>
<td>22</td>
<td>12</td>
<td>6</td>
<td>54</td>
</tr>
</tbody>
</table>
4 ‘WTP Enables More Comprehensive Valuations of Benefits’

The QALY, viewed as a health-related utility measure, provides for only a limited perspective on preferences related to a health care programme. WTP is ‘superior’, from a sovereign consumer perspective, in that it imposes no restrictions on which dimensions of benefit (health or otherwise), or for which attributes of the health care programme, respondents express their preferences. Neither does the method require the restrictive assumptions on the utility structure as does the QALY-algorithm, in particular that of time independence of health state utility.

4.1 The QALY - A Restrictive Measure of ‘Utility’

QALYs are normally considered to be a measure of utility which appears in the denominator of CUA. However, a closer look at the QALY algorithm shows that it includes important non-preference based parameters: $Q = (HS_1 - HS_0) \times T_D \times p$.

While $Q$ is the outcome, the only utility based parameters are those relating to the health states; $HS_1$ and $HS_0$. Preferences for the duration of the health improvement, $T_D$, may be accounted for if a preference based discount factor is used. Preferences related to the probability of successful outcome, $p$, could be accounted for if standard gamble were used as the elicitation-technique. A compound utility measure for the three terms (health states, time and probability) is offered by the healthy-year-equivalent, HYE (Mehrez and Gafni 1989).

However, still there are restrictions with $Q$ as a utility based outcome measure. First, there is no concern for preferences related to one’s own probability of ever needing the service (option value), nor for equity or distributional preferences (externalities) related to those patients who are to benefit. Second, there is a neglect of any disutility related to the actual diagnosis or intervention per se. Third, QALYs show no concern for programme specific preferences (process utility). Finally, individual variations in the intensities of health state preferences are neglected as every utility score is compressed/normalised within the same end-points of 0-1.

Both QALYs and WTP are instruments for expressing intensities in strength of preference. There is, however, a crucial difference between the cardinal scales. While health state utility scales are constrained by the same maximum value of 1 (reflecting, variously, ‘perfect health’ or ‘best imaginable health’) for each individual, WTP scales do not have such a uniform upper limit. It is constrained by the respondent’s budget, which of course will vary among income groups. Hence, as opposed to QALYs, WTP is not an income-neutral preference scale.

In judging the extent to which WTP in practice offers a more comprehensive utility measure we would have to inquire into how many of the above characteristics are included in the scenario descriptions which have been presented for subjects when they have been asked to express their WTP. As explained above, there are three essential dimensions of health outcomes; health state improvement, duration and probability. When applying WTP as a benefit measure, this outcome can be described in a more comprehensive way in terms of a health gain profile, as opposed to the QALY approach which is based on partial valuations of health states only.
4.2 Beyond Health Outcomes

It is often seen as an argument for WTP that "CBA has a broader scope than CEA/CEA" (Drummond et al 1997). As expressed by Donaldson et al (1997): "The advantage of WTP over QALYs stems from the fact that the latter permits the valuation of health gains only". What WTP is capable of doing in theory is to account for peoples’ valuation of other characteristics of health care beyond health outcomes.

Mooney (1991) has argued that, beyond QALYs, health care may ‘produce’ non-health benefits such as information, caring, anxiety reduction and communication. Further, he argues that there is utility from the ‘process’ by which care is provided (e.g. degree of patient autonomy), which he calls ‘process utility’. Smith (1996) has also argued for the existence of ‘regret’ in the value function of individuals expressing preferences over health interventions. Some of the empirical surveys are very explicitly using WTP as a measure of valuing such non-outcome benefits (e.g. Ryan 1996, Donaldson et al 1997, Donaldson and Shackley 1997).

An inquiry into what lies beyond health outcomes must first clarify what lies within (Olsen 1996). Although an inclusive outcome measure may be used, there is still the argument that there are more ‘utility affecting characteristics’ of a programme than those which can be captured in terms of QALYs. The scenario specification, the description which respondents have been asked to value, is therefore crucial (Smith and Dobson, 1993). A prerequisite for a good WTP-survey is to provide a careful and understandable description of the characteristics of the programme which are assumed to yield utility – both the health outcome dimensions and the non-outcome attributes. However, beyond these benefits from health care use there are potential benefits in terms of option value and externalities.

4.3 WTP: Use Value, Option Value and Externalities

In the environmental economics literature, a distinction is made on which aspects of the good people value. For environmental goods, they are: (i) use value; (ii) option value; and (iii) existence, or non-use, value. The parallel to health care is use value for current patients, and option value for the possibility of needing care in the future (i.e. the insurance motive for being willing to pay). However, rather than existence value, the third type of value for which people may express a positive WTP lies in the caring externality for other peoples’ health (care use), expressed as a subsidy motive, or a more selfishly explained externality in terms of reduced contagion. The distinction between these three types of values was also emphasized by O’Brien and Gafni (1996), who distinguished between three types of respondents to a WTP survey; (i) currently diseased, (ii) currently non-diseased but at future risk, and (iii) currently non-diseased and not at future risk.

The perspective of the analysis, that is the role in which subjects are being placed, determines which values are being activated. If current patients are asked, we would expect use value to be the predominant (although not exclusive) type of value expressed. That is the ex post perspective undertaken in those studies which have asked patients, or target groups with an identified need for the intervention in question. There are two different ex ante perspectives for subjects who could be at future risk. The first is a private insurance premium which would express the option value. The nature of private insurance is to have the premium reflect one’s own risk, hence, there is no element of ex ante cross subsidisation to express caring externalities. The second is a social insurance in the form of a tax
contribution, which would express a combination of an option value for oneself plus a possible caring externality for others. Finally, subjects could be asked to express their WTP solely for other people’s use of care, i.e. a caring externality. If a representative sample of the general population is asked, that would include all motivations for WTP; use value, option value, and caring externality. A comprehensive valuation of benefits should ensure that all these valuations are being accounted for.

When reviewing the scope of valuations conducted in WTP surveys, we would therefore inquire into who has been asked and from which perspective. Because WTP has the potential to value more comprehensively all these analytically different sources of benefit; i.e. use value, option value and externalities, we were concerned with the extent to which the surveys had actually exhausted its potential on this front. For each survey we carefully assessed whether it was likely that respondents would account for any option value and/or externalities. Three types of respondents have been asked; users, convenient samples, and the general population. The users are most often patients with an identified need for the treatment, or sometimes their relatives. Convenient samples are most often groups of students or seminar participants. General population samples are those who are supposed to be representative of the population. Table 2 shows that most surveys included only use value, reflecting that most surveys asked patients only.

Table 2: Different Values by Type of Respondents

<table>
<thead>
<tr>
<th>RESPTYPE</th>
<th>Use Value</th>
<th>Option Value</th>
<th>Extern.</th>
<th>Option &amp; Extern.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>34</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>Conven sample</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>General pop</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>User+Conven</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>User+Genpop</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Conven+Genpop</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>41</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>54</td>
</tr>
</tbody>
</table>

Note that the vast majority of the surveys have asked users, and that only nine studies have asked the general population. When we suggest that users may have expressed option value or externalities, that would refer to preventive programmes (including vaccinations).

4.4 How Have They Been Asked?

There seems to be wide agreements in the literature that face-to-face interviewing is the much preferred method for obtaining reliable answers (NOAA 1993, Mitchell and Carson, 1989). Table 3 shows that this method for data collection is used only in 19 (35%) of the surveys. Self administered questionnaires handed out to users is a frequently chosen approach. Note that there are only 4 (!) studies which have collected WTP-data by face-to-face interviewing from the general population.
Table 3  Data Collection by Type of Respondents

<table>
<thead>
<tr>
<th>RESPTYPE</th>
<th>Face2face</th>
<th>Telephone</th>
<th>Postal</th>
<th>Self Adm</th>
<th>Not Available</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>12</td>
<td>4</td>
<td>9</td>
<td>12</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>Conven sample</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>General pop</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>User+Conven</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>User+Genpop</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Conven+Genpop</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>19</td>
<td>6</td>
<td>12</td>
<td>16</td>
<td>1</td>
<td>54</td>
</tr>
</tbody>
</table>

4.5 Valuing What?

Outcome descriptions were originally reviewed from the criterion of how many of the three dimensions of health outcomes were included in the scenario; health states, duration, probability. Only 16 surveys had any notion of risk or probability explicitly pointed out to subjects. If at least two dimensions were included, we labelled it a 'comprehensive description'. As few of them passed this criterion, we introduced the notion of 'understandable description' for surveys whose outcome descriptions were believed to have been meaningful to the respondents, due primarily to being patients familiar with the diagnosis. If respondents must have been thought to understand what the intervention would mean in terms of outcomes, we labelled them 'implicitly thought to be understood'. Lastly, some studies were explicitly not concerned with the valuation of health outcomes, but rather with non-health aspects of the intervention. These were labelled 'not the focus'. Some papers gave no indication of what had been presented to respondents, neither the questions asked nor the scenarios presented – they have been labelled 'not available'.

Table 4: Degrees of Comprehensive Outcome Descriptions

<table>
<thead>
<tr>
<th>Degrees of Description</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensive</td>
<td>11</td>
<td>20,4</td>
</tr>
<tr>
<td>Understandable</td>
<td>15</td>
<td>27,8</td>
</tr>
<tr>
<td>Implicit</td>
<td>12</td>
<td>22,2</td>
</tr>
<tr>
<td>Not the Focus</td>
<td>12</td>
<td>22,2</td>
</tr>
<tr>
<td>Not available</td>
<td>4</td>
<td>7,4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>54</td>
<td>100,0</td>
</tr>
</tbody>
</table>

Few surveys provided a comprehensive description of outcomes. With our best will, we found only 4 studies, whose description could facilitate any estimation of an implicit health outcome in QALY or HYE terms.
Some studies have focused very explicitly on valuations of non-outcomes attributes. Still, the scenario descriptions are equally important, in terms of what is supposed to be utility yielding non-outcome characteristics.

**Table 5: Degrees of Comprehensive Description of Non-Outcome Attributes**

<table>
<thead>
<tr>
<th>Degrees of Description</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensive/understandable</td>
<td>12</td>
<td>22.2</td>
</tr>
<tr>
<td>Implicit</td>
<td>15</td>
<td>27.8</td>
</tr>
<tr>
<td>Not the focus</td>
<td>26</td>
<td>48.1</td>
</tr>
<tr>
<td>Not available</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>54</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Among the 28 (54 – 26) studies which have a focus on non-outcome descriptions, only 12 are comprehensive (and understandable). However, among the 12 papers with an explicit focus on non-outcome characteristics (i.e. health outcomes were not the focus), 10 offered a comprehensive scenario description.

In which way was the scenario presented to the subjects? Only 26 had given a separate scenario description prior to the WTP question asked. The remaining would have had the good to be valued presented as part of the questionnaire, most often in terms of only mentioning the diagnosis. Our impression is that for the majority of studies there has been a WTP for a health care product rather than the health produced by the product.

Lastly on what have been valued, Table 6 shows the main diagnosis and the pharmacoeconomic relevance. There is a concentration on two types; cardiovascular diseases (e.g. hypertension treatments) and obstetrics/maternity care (e.g. IVF). There are fewer pharmacoeconomic studies than one might expect.

**Table 6: Main diagnoses by Pharmacoeconomic Relevance**

<table>
<thead>
<tr>
<th>DIAGNOSI</th>
<th>Explicit drugs</th>
<th>Implicit drugs</th>
<th>Non-drug</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovasc</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Obstet/maternity</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Primary/elderly</td>
<td>5</td>
<td>3</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Risk</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>More than one</td>
<td>1</td>
<td>2</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>9</td>
<td>15</td>
<td>30</td>
<td>54</td>
</tr>
</tbody>
</table>
5 ‘WTP can Advise on Overall Social Efficiency’

CEA/CUA cannot provide any guidance to decision makers in assessing the social optimality of a programme, i.e. whether resources should be spent on the programme at all, and if so, at what level. It is only when compared with C/E-ratios of alternative programmes that analysts could suggest that the new programme be implemented, by deferring a programme with a higher C/E-ratio. CEA/CUAs are unable to provide advice concerning the acceptable cut-off ratio, as the unit of outcomes is not commensurable with costs, nor with outcomes in other sectors of the economy. This is not merely a practical issue of the current level of development of these techniques. Even if CEA/CUA were developed to perfection, there would still be the theoretical issue of the degree to which they can assist in decisions concerning social efficiency.

Therefore, theoretically there is a compelling argument in favour of applying an outcome measure which is commensurable with the cost side, as well as with outcome measures in other sectors. In theory, a CBA enables analysts to give recommendation as to whether the project should be implemented: when it increases social welfare (positive net present value), and when there is no better alternative programme in any other sector of the economy. This is the line of reasoning in favour of CBA over CEA/CUA as presented in Drummond et al (1997). How powerful is this argument in the real world of public policy decision making?

Consider, for example, a situation in which government is faced with a CBA which has reported that the WTP for a new drug exceeds its marginal social costs. Would the government be able to conclude that this new drug should be funded by deferring a marginal programme from another public sector. It is not possible to determine whether the forgone benefits of the deferred programme are any less than the potential benefits from the new drug. The critical point is that: only in the general equilibrium situation, where marginal benefits equal marginal costs for all existing programmes, would such a theoretical argument hold. As long as we do not know that the deferred programmes are valued less than the new one, the analyst cannot infer that adding a new programme into a given total budget will increase social welfare.

In judging the relevance of WTP-surveys in aiding decision making, the important issues are which decision making body, and which type of choice, is WTP meant to assist? The simplest choice is a demand study for a new intervention (e.g. drug): how much could patients be WTP for this new drug? The context is a ‘market test’ of how much users would be WTP, with no separation of users from third party payers. Alternatively, the context could be how much patients could be WTP for a given elective surgery done by their preferred doctor and within a specific waiting time. This could be put in an ex ante context of how much one would be WTP for an insurance which were to cover the elective surgery with the same characteristics. Again, there is no separation of users and payers, since patients would either pay out of pocket or pay through private insurance. Furthermore, in principle, there would be no opportunity costs in terms of displaced health care provided to others.

Things, however, differ when the context is that of assisting a public health service in improving its allocative efficiency. Quite often health policy recommendations are given resulting from this, that, either explicitly concluded or implicitly inferred, the public health service should fund it. By definition, respondents have expressed that they are willing to sacrifice the alternative private consumption which their WTP could have purchased. Interestingly, the conclusion of this is that the only justifiable source of funding is therefore
private income! Hence, if this is not the actual source of funding, one has to make sure that the opportunity costs in terms of benefits forgone in the deferred public programme(s) are less than WTP. This we cannot tell without valuing the candidates to be deferred. Hence, if money is to be taken from a given health care budget, a partial valuation is of no aid to decision makers.

However, rather than being part of a CBA, WTP could alternatively be used as an instrument for expressing intensities of preferences and aid policy makers who consider two or more programmes, but do not know which the community may value most. The WTP survey could then provide policy advice on which of the alternative programmes is most valuable. Similarly would WTP surveys provide useful health policy information in situations where alternative interventions exist for one diagnosis, e.g. which of two close substitutes (with identical costs) is the most valuable.

5.1 WTP in Practice – Partiality, Implications And Mode Of Payment

Studies have overwhelmingly been partial evaluation, of a programme in isolation with no comparator. However, from Table 7 it seems that those studies which are potentially most helpful to decision makers in a public health service, i.e. those valuing close substitutes and those which compare the benefits from alternative resource uses, have increased most over time.

Table 7: Partiality by publication period
PUBPER * PARTIALI Cross-tabulation

<table>
<thead>
<tr>
<th>PUBPER</th>
<th>PARTIALI</th>
<th>One service</th>
<th>Close substitutes</th>
<th>Diff diagnosis</th>
<th>More than one service</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985-89</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>1990-93</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>1994-96</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>1997-98</td>
<td>12</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>TOTAL</td>
<td>38</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td></td>
<td>54</td>
</tr>
</tbody>
</table>

The implications of, or the lessons drawn from, the surveys were classified depending on whether the focus of the papers were primarily on methodological issues, primarily health policy, or both. Table 8 compares implications by type of journal (see Appendix 2 for our coding of journals).
Unsurprisingly, economics journals published the methodologically oriented papers, while clinical journals published surveys with a health policy focus.

The use of WTP may essentially serve two policy purposes; either it is a measure for variations in strength of preferences for alternative resource uses (including 'close substitutes'), or, it is a way of estimating the benefit side of a CBA in order to derive the net present value (or the cost/benefit ratio) of the programme in question. Only 16 papers compared the WTP with the cost of the programmes and suggested the magnitude of the net present value. Interestingly, Table 9 shows that those studies which did so had a more policy focus. The methodological papers were essentially reporting WTP only.

Table 9: Implication by CBA-Measure
IMPLICIT * CBAMEASU Cross-tabulation

<table>
<thead>
<tr>
<th>CBAMEASU</th>
<th>WTP only</th>
<th>NPV</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMPLICIT Methodological</td>
<td>19</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>Health policy</td>
<td>11</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Both</td>
<td>8</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>TOTAL</td>
<td>38</td>
<td>16</td>
<td>54</td>
</tr>
<tr>
<td>Mode of Payment</td>
<td>Frequency</td>
<td>Percent</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Out of Pocket (oop)</td>
<td>42</td>
<td>77.8</td>
<td></td>
</tr>
<tr>
<td>Tax</td>
<td>4</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>Oop+tax</td>
<td>1</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Oop+insurance</td>
<td>1</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Tax+voluntary</td>
<td>1</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Oop+tax+insur</td>
<td>1</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Not available</td>
<td>4</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>54</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
6 Two Health Policy Issues: Equity and Transferability of Results

6.1 Equity

WTP involves complete suppression of distributional considerations. The *social* value of a programme is measured through a simple summation of all individuals’ expressed preferences assessed in terms of WTP, i.e. the premise of ‘one dollar one vote’, and that ‘a dollar is a dollar is a dollar’ (analogous to ‘a QALY is a QALY is a QALY’ in CUA). The issue here becomes a dichotomous one of whether: (i) income is a significant determining factor of WTP; and (ii) if so, what measure can be taken to ‘correct this bias’. In theory at least, as long as WTP depends on ATP (ability to pay) those with a greater ATP will, *ceterus paribus*, be able to express a greater WTP. The equity implications of this are simple: if one wealthy individual is willing to pay $100 for programme A and 10 less wealthy individuals are willing to pay $10 each for programme B, the two programmes would be perceived to yield the same social welfare, quite independent of any differences in the aggregated health produced. Thus, the determining factor over the provision of programme A or B would be the relative income of the two groups. In practice the issue is the degree to which the differential income constraint would cause such a ‘skew’ in WTP values, which is likely to be significant. Furthermore, there is a large body of evidence that ‘needs’ (in terms of severity or capacity to benefit) differ between income groups. In this case using WTP unadjusted for income will skew resource allocation to the preferences expressed by the wealthy.

Table 11 shows that only 11 studies gave explicit recognition to the potential inequitable implications of basing resource allocation on preferences as expressed through WTP. Interestingly it was in the early days that equity aspects received relatively most concern. It should be noted, though, that among the seven studies published since 1990 which have given explicit recognition to distribution, five are by the same author. Had it not been for him, explicit equity considerations would seem to have vanished completely in these studies.

Table 11: Recognition of (inequitable) distribution by publication period

<table>
<thead>
<tr>
<th>PUBPER</th>
<th>Recognition of Distribution</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No mentioning</td>
<td>Explicit recognition</td>
</tr>
<tr>
<td>1985-89 (5 years)</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1990-93 (4 years)</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>1994-96 (3 years)</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>1997-98 (1.5 years)</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>43</td>
<td>11</td>
</tr>
</tbody>
</table>

Interestingly, over time there has been a different handling of the income variable than as a concern for inequitable policy. The minimum validity testing undertaken appear to be an inquiry into any correlation of income with WTP (25 studies). If significant, that is taken as a measure of ‘theoretical validity’. It appears somewhat peculiar that the discovery of a correlation between income and WTP is taken as an indication of the validity of the *absolute level* of the WTP values.
6.2 Limited Transferability of Outcome Values

The idea with QALYs was to develop a generic outcome measure which is transferable across programmes. In so doing, the focus of value is on the selected dimensions of health and their various levels. The health outcomes from alternative programmes are compared by how they score on the different dimensions (in the way of a multi-attribute-utility function). Following this procedure, preferences for other characteristics of the programme which are not being accounted for in these predefined levels and dimensions are ignored.

The application of an outcome measure which is preference inclusive (WTP rather than QALYs) involves that all attributes specific to this programme or intervention are being valued in aggregate. The implication is of course that the aggregated WTP-value cannot be transferred to the valuation of programmes which have different combinations of characteristics. When no attempts are made to attach WTP values to each level of improvement on each characteristic, the informational value of the survey is indeed restricted to the specific programme. If no transferability, a separate WTP survey has to be undertaken for each possible programme, which of course is very costly. Interestingly, we could not find any study which suggested that the obtained WTP figures could be transferred to the valuation of other programmes.

It is sometimes held that WTP surveys are most appropriate for avoiding short term unpleasant events, while QALYs are appropriate for long term effects. While this may be true (at least it offers an appealing harmonious idea of complementarity rather than substitutability), the resource allocation implication is still that we cannot arrive at an optimal split of the budget between the two areas.
Conclusion

Our reading of the literature tells that there are three types of arguments used in favour of applying WTP in health care. The first; "WTP is theoretically correct because of its foundation in welfare economics, therefore it should be used to value health care", is indeed no argument per se. At best it is a normative contention which may be discussed elsewhere.

The second that "WTP enables more comprehensive valuation of benefits than QALYs" is intuitively more appealing. In principle there are three sets of benefits from health care; 1) health outcomes, measured by the three key dimensions; health state, duration and probability, 2) beyond health outcomes, in terms of utility yielding characteristics of the intervention, and 3) non-use values, i.e. option value and externalities. We have reviewed the empirical studies with a focus on how the benefits were described to respondents.

Accounting for a comprehensive description of benefits is cognitively demanding. It is therefore required with a separate scenario description prior to the valuation exercise (done in 26 studies). In judging the comprehensiveness of the scenarios, our review shows that non-use values were rarely accounted for. Those studies which had an explicit focus on non-outcome characteristics had given a comprehensive description of these characteristics. The descriptions of health outcomes appear quite incomplete. In only four surveys could we make any estimation of health outcomes in QALY-terms. Hence, there seems to be little scope for using the existing studies as basis for deriving monetary values on health outcomes.

The third argument is the CBA-rationale of improving allocative efficiency. We have taken the view that it is impossible to infer from a partial WTP study (38 studies were partial) that a new programme should be funded from a given total health care budget. This we cannot tell without knowing the opportunity costs in terms of the benefits forgone from the displaced programmes.


APPENDIX 1

The Complete Reference List of Papers Reviewed


treatment in obese men with mild hypertension’, *Journal of Hypertension*, vol 10, no 9, pp
1063-1070.


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.


Kartman B, Andersson F, Johannesson M 1996, ‘Willingness to pay for reductions in angina

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.


information on patients’ willingness to pay for autologous blood donation’, *Medical Care* (in
press).


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


willingness-to-pay approach’, *PharmacoEconomics*, vol 8, no 1, pp 34-45.

pharmaceutical: A feasibility study of contingent valuation in managed care’, *Medical Care*,
(forthcoming).


Classifying into Main Types of Journals

Clinical journals (14)
- J Pediatrics (2)
- Arch Fam Plan (1)
- Birth (1)
- Acta Obstet Gyn Scan (1)
- Contraception (1)
- J Clinical Pharmacy (1)
- Epidemi Infect (1)
- J Internal Medicine (1)
- A Allergy Astma Infect (1)
- J Advanced Nursing (1)
- J Hypertension (1)
- Drug Intell Clin Pharm (1)
- British J Urology (1)

Public health journals (22)
- Soc Sci & Med (5)
- Medical Care (5)
- Med Decision Making (4)
- Int J Tecn Assess HC (2)
- Health Policy (1)
- J Clin Epidemiology (1)
- Am J Public Health (1)
- Health Care Analysis (1)
- Can J Public Health (1)
- Leadership in HC (1)

Health economics journals (12)
- Health Economics (5)
- J Health Economics (4)
- Pharmacoeconomics (1)
- J Res Pharm Econ (1)
- Presentation at a health economist meeting, yet unpublished (1)

Non-health economics journals (6)
- Applied Economics (3)
- J Public Economics (1)
- J Social Economics (1)
- J Risk &Uncertainty (1)