The willingness-to-pay concept in question

ABSTRACT

The adequacy of the concept of willingness to pay within health economics evaluations is reviewed. A considerable number of researchers in the literature have pointed out multiple methodological issues involving willingness-to-pay estimates. On the other hand, the theoretical discussion about the aggregation of individual preferences within an aggregate demand remains open. However, over the last 20 years, willingness-to-pay estimates alongside health economics research significantly increased and in many cases they are one of the key factors for decision making on issues of health policies. The article describes some limitations of this approach as well as the potential distorting effect that it might have on health economics evaluations.

INTRODUCTION

The concept of willingness-to-pay (WTP) has become very popular over the last twenty years in economic assessment studies in the health field.25 WTP is a methodological tool that seeks to estimate the capacity to pay of certain social groups in a search to find out the hypothetical monetary value for programs and specific medical interventions and treatments. Its application in cost benefit assessments and in decision-making processes in other economic assessment models have made this tool one of the most requested in the area of health economics. Nevertheless, there is still much to analyze in this topic since its use is debatable in some cases.39 For example, some authors base their concern on the way in which information is presented in the questionnaires (its order and characteristics), which have proved to be decisive to establish the results of subsequent WTP estimations.29,34-37 On the other hand, not only could this type of methodological problem skew the WTP estimate, but the direct and indirect experience of those surveyed about the treatment being analyzed could this type of methodological problem skew the WTP estimate.29,34-37

Despite the methodological problems, other conceptual problems could be taken into account in deciding about the use of this technique.24 One of them is the discussion about whether individual preferences can be aggregated within a social function.15 Similarly, the less-than-perfect aspects of the instruments that are used for achieving a good approximation to reality reduces their internal and external validity, thus generating serious problems as to the reliability and credibility of the measurements.

THE CONCEPT OF WTP IN THE HEALTH FIELD

The cost-benefit analysis compares the discounted future benefit flows of a particular program with its cost flow; the difference between the two corresponds to the net social benefit of the program.16 Considering that the cost-benefit analysis estimates the costs and the benefits of health in monetary terms, some authors use WTP (even though there are problems with it) to inform their decisions within the different sectors of the economy. Even so, cost-benefit analysis has the advantage of capturing the positive and negative effects that are outside the application area and that are known in economics as externalities.16

WTP is a survey method that presents the interviewees with hypothetical scenarios about a certain intervention or specific program which it is intended to evaluate. Based on a real market for a specific health program or benefit, the survey consults participants as to the maximum they would be prepared to pay for such a service.6,16 Within the health field the vast majority of the investigations that use this methodology call them contingent valuation studies,4 and they try and consider the worth of something that is not on the market, for example a potential health program, the true value of which is estimated by means of collective financing. It is for this reason that in contingent valuation studies real or potential consumers are asked to consider how much of their income they would be prepared to sacrifice to have the benefits of a private health plan, if they were available in the market.

The concept of contingent market is powerful and can be used to assign monetary values to all the suggested elements of a health program and not simply to the value of health. This is relevant for all decision-makers who could use it as a broad measure of health results.33 Drummond et al16 point out that it is important to remember that this is a very variable measurement technique and one that depends on the structure and design of the questionnaires. This technique only gathers the perceptions and preferences of those surveyed about a hypothetical scenario that is linked to the different alternatives of medical intervention and the prior information that the interviewee has. However, there is currently significant disagreement in the academic world about how to measure the WTP of a society.33

Interesting methodological investigations have concentrated their efforts on avoiding the potential bias that may be produced when estimating WTP. Apparently, the estimates that usually contain a smaller bias are those WTP surveys that use open questions, since the pressure on the person surveyed to come up with a monetary response is less. This type of design, however, generates greater imprecision with very variable replies or an absence of replies altogether (people who do not want to reveal their preferences).26 Equally, literature usually indicates that many surveys with closed questions do not obtain conclusive results,38 particularly for dichotomous questions.6 In many cases, it is preferable to use questions with some type of scale (Likert scale).

As was previously mentioned, the WTP measurement technique proposes the application of a questionnaire to measure willingness to pay (correspondents in a first measurement round). Considering that the results might be potentially skewed, the application of various techniques has been put forward in an attempt to correct them and to bring them closer to the true values.
Among these we find the carrying out of a second round for obtaining a WTP, after a brief explanation of the results and their impact,\(^5\,^6\,^7\,^8\) and even a third round where selection of the WTP results is adjusted according to the degree of certainty of buying and/or acquiring the program. In these cases, only the WTP results obtained from buyers who are extremely likely to make the purchase are used.\(^9\,^10\,^11\) Otherwise, the WTP values would be overestimated.\(^12\) In this way its level of bias is reduced, albeit not fully corrected.\(^13\)

On the other hand, because there is no true market with which to compare the results encountered in WTP estimates, establishing a validity criterion becomes almost impossible.\(^14\)

**LIMITATIONS IN THE WTP CONCEPT**

The WTP concept is clearly deficient for application in the health field. Ten reasons are mentioned by various authors.\(^15\,^16\,^17\,^18\,^19\) These are both theoretical and methodological and lead to questioning about the use of this tool:

1. WTP is a variable indicator that depends on the economic and social stratum in which the survey is carried out. Evidently, this has a strong impact when decision-making is based on a WTP measurement. For example, if only a high socioeconomic stratum is considered, it might be that a large number of cost-effectiveness studies that evaluate treatment with different medications or health programs are accepted as cost-effective if the decision variable is precisely willingness to pay. If the same studies only consider low socioeconomic strata, the results might have exactly the opposite results to the first study. This problem may lead to the fact that certain alternatives, which are proven to be efficient and effective, may be reserved for people according to their capacity to pay (a moral problem). One example is the use of antiretroviral drugs for treating HIV/AIDS; if the decision-variable is WTP this would result in them being cost-effective in high income societies (like Europe and North America) and not very cost-effective in other societies with lower incomes, like Africa. Obviously, since this deals with the same medication, how is it possible that the treatment is more convenient in some societies than in others, depending on their capacity to pay? The first limitation of WTP indicates that this instrument is clearly discriminatory for economic and not health reasons. From this point of view, few costly medical interventions should be applied to low income individuals and social groups when health is a right in the vast majority of countries and not something that can be traded like in any other market. To a certain extent, this fact might lead us to suppose that in certain societies the concept of WTP should not be recommended for decision-making if no care is taken when estimating it.

2. Nobel Prize winner for Economics, Kenneth J. Arrow, in a highly important and influential article, showed how impossible it is to extrapolate individual preferences within a social function.\(^20\) In this sense, what is the theoretical and empirical justification for using measures like WTP? The assumptions used by Arrow included its efficiency in the Pareto sense, independence of preferences and that these should not be dictated upon (other than the assumption that preferences must be complete and transitive). On this point, economics literature has had an extensive debate and the aggregation of individual preferences are generally accepted when the convexity assumption and interpersonal comparison of such preferences is imposed.\(^21\) However, the Pareto criterion is still the most questionable point of the whole discussion. The Pareto efficiency is based on the fact that if an order of usefulness is established at the individual level \(a > a_1\), then in a social environment the same inequality should continue applying, when this may not be right.\(^22\) For the time being, the controversy still remains in literature, as does whether the aggregation of individual preferences is correct or not.

3. Another problem linked to the instruments of WTP is due to the significant information asymmetries that exist between patients and medical practice. For example, within the preferences the patients express the theoretical and practical knowledge that they may have with regard to the efficiency and effectiveness of various medications or interventions is questionable. Even when the initial or prior information is sufficiently good, the guidance that might be given to it is debatable. Subsequently, generating a social function based on few or badly informed individual preferences may lead to questionable results. A demonstration of this latter example was found in Norway\(^23\) where the interviewees gave the same implementation cost to three different health programs (a helicopter ambulance service, an increase in the budget for heart operations and an increase in the budget for hip operations). In that particular study, the investigators showed that the use of WTP was feasible for some medical interventions, but the conclusion was that it was necessary to develop this method better to obtain conclusive results.

Some WTP studies have tried to solve this problem by asking medical personnel about WTP. Even when in these cases the problems of lack of clinical knowledge are reduced significantly, an attempt is generally made to have the opinions of the doctors be representative of a whole social group or a national health system, which generates a conceptual problem when it overlaps individual preferences and social needs or the needs of the public health services. An example of this latter occurred in an investigation in Canada where the WTP for Docetaxel in treating advanced ovarian cancer was estimated based on the replies of 80 oncologists and nurses, and regarded as being representative of the
country’s health system.\textsuperscript{14} The following question then arises: Is it conceptually valid and representative for 80 health workers to be the decisive factor in a global health system? Do they really show the WTP of an institution?

4. Another bias problem mentioned in literature has to do with the fact that many interviewees concentrate more on the costs of medical intervention rather than on its results, which is why they are not valued as is required.\textsuperscript{3} Even when these problems have been partially solved by using the marginal approach methodology, a lot still needs to be developed in this sense.\textsuperscript{29,30} For example, although this mechanism reduces preference reversals, it does not allow for no comparison with all the health services that are of interest to an individual or society (questions are only asked about comparable treatments) and so the results for WTP are maximized or overestimated.\textsuperscript{29} Equally, this new methodology still does not produce consistent results within the replies.\textsuperscript{35}

5. Even so, WTP estimation calculations might be overestimated from the social point of view in those studies where the instruments are applied only to patients and the families of patients affected by a particular illness and not to society as a whole (which evidently would provide a lower valuation relative to the former).

This last fact would tend to raise WTP estimates for a particular society and accept some medical interventions as cost-effective when, in reality, they are not. Examples are cases where a question is asked regarding what percentage of a patient’s or family’s income would earmark for certain treatments according to the degree of efficiency or effectiveness of a new drug. This point is very much discussed, because some investigators indicate that after applying the questionnaire the next step is to extract an aggregate demand curve,\textsuperscript{7} which depends significantly on the way in which the questions are worded. Even so their validity is debatable in society since one is not talking of real demand but suggested (hypothetical) demand. This discussion on placing hypothetical preferences over real ones is an old debate which has been going on in economic literature for more than 60 years.\textsuperscript{8} In this sense, it is suggested that the WTP concept be used within the context of private insurance rather than an institutional context, in which the participation of those being surveyed may have a potential representativeness over hypothetical demand.\textsuperscript{2,6}

In the literature of health economics, numerous attempts have been made to find close connection between the hypothetical WTP and the actual WTP.\textsuperscript{41} Results have not been conclusive. For example, in an asthma treatment study carried out in 2001 in the United States it was determined that the difference between the hypothetical WTP and the actual WTP was significant (38\% vs. 12\%; p<0.001), showing a clear overestimation of potential demand.\textsuperscript{8} In this sense, it has been shown that, even empirically, there are difficulties with this tool when it comes to inferring real social demand for new treatment or new technology. Generally speaking, as economic literature indicates, these types of treatment are overestimated.\textsuperscript{20} This has been proved in different meta-analyses\textsuperscript{18,25} even when techniques are applied for correcting estimation bias, it is reduced but not eliminated.

6. On the other hand, it is possible that WTP estimates are underestimated since willingness to pay may be insufficient to cover a predetermined optimal service from public health institutions. This supports the argument that health is a right of all citizens and that even when the population’s capacity to pay is limited, some medical services have to be provided for political, ethical and equity reasons. In other words, the WTP concept only takes into consideration the individual willingness to pay of a society and excludes, in the vast majority of cases, the collective (social) capacity. For example, these analyses usually forget there are governmental and non-governmental organizations (NGOs) prepared to pay huge amounts of money to contain an epidemic or reduce morbidity rates in low income regions. The best example of this is the vaccination campaigns on the outskirts of cities or in rural areas in developing countries where the capacity to pay is minimal and where national health systems provide services that produce very favorable results.

7. WTP estimations might show that there are methodological problems of the representativeness of the sample (problems of external validity, size of the sample, with over- and under-represented groups, in accordance with socioeconomic, age and gender strata) or some other type of bias, in which questions are prepared in such a way as to induce a certain type of response. Even though today there are already some methodological guides that are useful for minimizing problems with bias, it is still relatively frequent. With regard to this last point, some authors have indicated the different consequences of the order of the questions in WTP questionnaires.\textsuperscript{23,39} One of them is that a conceptual problem arises, since the economic rational choice theory assumes that in order to establish a consistent order it has to be maintained regardless of the sequence in which the different alternatives are presented.\textsuperscript{22} Empirically, there is evidence that this happens. A study in Ireland showed that the presentation order of different treatments had a strong impact on the monetary estimates that the interviewees assigned to them.\textsuperscript{39} This obviously skewed the investigation in favor of the first options in detriment to the last ones. Regardless of the results of the Irish investigation, the use of conjoint analysis in studies has proved that the order of the question is of little importance to the answers.\textsuperscript{32} Some investigations have indicated that, methodologically, these models are both extremely
valid and internally consistent. However, even with this methodology, the discussion about the external validity of the investigations and the use of preferences for generating a social function of aggregated demand is still open to question. For example, various authors have pointed out the statistical problems associated with obtaining the WTP average in cases of non-parametric distribution. However, such limitations have been largely solved.11,12,22,43

Finally, the psychological and technical aspects of WTP studies must be added.9 The first deals with the difficulty of solving the problem associated with questionnaires that are used to collect the information that generate a low response rate. The second has to do with the robustness of the statistical analysis of these studies. The problem comes from WTP questionnaires that present a very small number of discrete selection responses relative to the number of scenarios generated. This evidently complicates the resulting statistical analysis when this technique is used.

8. On the other hand, health market prices are often not established in accordance with competitive parameters (but rather they behave in a monopolistic or oligopolistic way) which is why they reflect no type of scarcity. Therefore, WTP does not always suggest the advantage for a society to adopt one treatment rather than another, since some market costs are distorted by non-competitive factors (for example, the monopolistic price of drugs in some cases).

However, in health economics literature it is usually stated that governments take decisions with the purpose of maximizing social well-being, especially when there are competitive markets.10 For this reason, when there are no competitive markets (as is the case in the health sector) the cost-benefit analysis technique is suggested, which includes an estimation of WTP. Nevertheless, most of these investigations still do not differentiate the economic factors that might be distorting the market, centering the investigation on a hypothetical aggregated demand when whoever defines the price and the amount offered of a particular medication or medical intervention in the health sector generally comes from the supply side. The only exception to this are those drugs that are available because of international or local agreements between pharmaceutical companies and national health systems. From this perspective society’s WTP plays a minor role in the provision of new medical care services.

9. WTP studies for various countries or regions are not comparable since preferences are different. Generally speaking, preferences derive from the individual’s own circumstances, which include education, culture, income, health, environment, among others. Most of the time, the factors that affect preferences vary from one region or country to another, which is why multi-regional preferences do not start from the same scale nor do they have the same cultural pattern, especially on themes like health. For this reason, it is not feasible to use this type of study for comparison purposes, or for drawing aggregated conclusions. Nevertheless, in Europe, a prominent attempt has been made (EuroWill Project) in six countries to determine the feasibility of the contingent valuation method as a tool for measuring the preferences of the general public with regard to different health programs. In short, this is something that even those in favor of this technique must study in more depth: generalization of the results.9

10. Finally, WTP is only useful for some perspectives of the investigation, for example, from the patient’s point of view. However, from other perspectives its use is highly debatable, as is the case with the public health service provider, where the provision of service is beyond the insured’s capacity to pay. Whether the beneficiaries of the public health service have high or low incomes, the services required must be provided regardless of the wishes or the capacity to pay of families. On the other hand, even when the beneficiaries of a particular health system wish to earmark more of their income for the treatment of an illness (greater WTP) given the current taxation model, this would be extremely complicated over the short term since tax collection for the health field is unique and indivisible (unless this increased expenditure is used for medications or other medical interventions in the private sector).

Another inconvenience that the WTP concept has in this sense is that individuals in most countries do not directly pay for medical care. This means that in many cases costs need to be deduced, which implies an administrative collection system that allows new medical treatment to be put into practice. So WTP estimates would be overestimated, since the person taking part in the survey, when replying to the questionnaire, does not notice facts such as these; he/she might only see the benefits of the intervention (and its direct costs) and not the indirect costs that interventions generate within the health system. For this reason, WTP is considered to be much more valid in a private context than a public one.

Despite what has been said, the controversy as to whether WTP really represents the ability to pay and the level of well-being of families, and if these terms are equivalent or not, still persists. Other studies mention that redistribution of funds within the home when there is spending on health, a rede...
CONCLUSION

The WTP concept is highly controversial in economic literature. On the one hand, its use has expanded enormously over the last 20 years in economic evaluations in the health area. On the other hand, the internal and external validity of such a tool is still questioned, both theoretically and methodologically.

It is even common today for different technology evaluation agencies to seek to take decisions based on predefined values, as is the case with the National Institute for Clinical Excellence in the United Kingdom for defining if a new intervention will be accepted or not (around £25,000-£30,000 per QALY gained). However, so far it is unknown where these values come from (their methodology and estimates) or whether they apply to all pathologies that exist in the health field. Interesting facts are currently being discovered about willingness to pay, such as that men are more willing to pay than women, or that there is no willingness to pay for long term medical care, such as was found in the investigation in Catalonia.

REFERENCES


33. Ryan M. Discrete choice experiments in Health care. BMJ. 2004;328(7436):360-1. DOI: 10.1136/bmj.328.7436.360


37. Smith RD. Contingent valuation in health care: does it matter how the “Good” is described. Health Econ. 2007;17(5):607-17. DOI: 10.1002/hec.1280


