Working Paper No. 2012/24

Pharmaceutical Pricing in a Globalized World: Crossing the income divide and ability to pay

Rutger P. Daems, PhD¹, Edith L. Maes, DBA²,
Christoph Glaetzer, Dipl. Kfm.³

September 2012

© The authors, 2012

¹ Planet Strategy Group, Brussels, Belgium
² Maastricht School of Management, PO Box 1203, 6201 BE Maastricht, The Netherlands; E-mail address: edith.maes@telenet.be
³ Janssen Asia Pacific, Singapore
The Maastricht School of Management is a leading provider of management education with worldwide presence. Our mission is to enhance the management capacity of professionals and organizations in and for emerging economies and developing countries with the objective to substantially contribute to the development of these societies.

www.msm.nl

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the School or its sponsors, of any of the views expressed.
ABSTRACT

OBJECTIVES

To examine pharmaceutical pricing strategies not only through the lens of efficiency but equity as well, and propose analytic methods that support policymakers and executives in making pricing and reimbursement decisions.

METHODS

The paper takes a 3-dimensional approach (triangulation) in defining international pricing policy for pharmaceuticals using: cost-effectiveness analysis (CEA), willingness to pay analysis (WTP), and ability to pay analysis (ATP). It attempts to find a balance between various economic methods of which some focus on effectiveness while others are geared towards incorporating equity in the equation.

RESULTS

A model has been developed to assess the ATP of 120 countries based on a country’s score in the human development index (HDI) as defined and measured by the United Nations Development Program, and published annually in the United Nations Human Development Report. The non-linear index has been applied to a number of drugs in different therapeutic categories without disclosing brand names. Prices in affluent countries – and to an increasing extent in the fast-growing middle-income countries – could generate sufficient revenue to pay for the cost of innovation benefiting all parties, whereas prices in the lowest-income countries in principle need only cover their marginal cost.

CONCLUSIONS

Besides the standard cost-effectiveness analysis (CEA) and willingness to pay analysis (WTP), a third method should be used in conjunction that measures the ability to pay (ATP) with the HDI index as yardstick. We recommend that ATP becomes an additional practice in policy decision-making and in defining international pricing strategies for pharmaceuticals in order to provide sustainable access to medicines.

KEYWORDS

Access-to-medicines, international pricing, cost-effectiveness, willingness-to-pay, ability-to-pay, equity.
INTRODUCTION

Providers of health technology are increasingly urged to give payers an insight into the associated cost and benefits, and demonstrate value for money. This necessitates the use of cost-effectiveness analysis in defining strategies that support pricing and reimbursement decisions. Although cost-effectiveness is an important criterion it does not provide a complete picture. A good health policy and related pricing strategy must not only aim to be efficient but should also ensure equitable access to medicines.

To fulfil the requirements of Health Technology Assessment (HTA) agencies and reimbursement committees in an increasing number of countries this requires not only generating data that prove clinical efficiency but also cost-effectiveness. The premise of cost-effectiveness analysis (CEA) is that it helps policy makers and executives make decisions by setting a maximum cost threshold for a benefit outcome, often expressed as quality-adjusted life year (QALY) gained. Reimbursement is granted if the incremental cost-effectiveness ratio (ICER) falls within an acceptable range where there is debate about its cut-off point. However, cost-effectiveness methods provide an incomplete picture especially when they are conducted from an institutional perspective. In cost-effectiveness analysis the efficiency-driven outcome is measured against a standard that reflects the economic considerations of the health system as a whole more than the willingness-to-pay (WTP) of patients and citizens.

This has led to a search on how to incorporate a patient’s or society’s willingness to pay for new and existing health technologies. WTP/QALY assessments are often based on contingent valuation techniques. Contingent valuation is based on a hypothetical market that uses the price an individual is willing to pay to obtain a beneficial intervention. However, the estimation of WTP is subject to considerable variability. The challenge is to avoid biases in surveying populations that due to government and insurance coverage may underestimate the preparedness to pay. The WTP/QALY ratios derived from specific patient populations may not accurately reflect the attitudes of society. The WTP/QALY from data describing human behaviour or preferences (utilities) not only yields disparate results, the process of aggregating utility outcomes is not firmly grounded in theory. Last but not least, the willingness to pay for improved health is influenced by income and standards of living.

Hence, a third method is proposed aimed at introducing proportionality by means of purchasing power parity measured at either the national or household level. The method is based on the human development index and deemed relevant especially for a global market confronted with disparity in the ability-to-pay (ATP) between countries and between population segments. This ATP method has been presented as a poster at the ISPOR 5th Asia-Pacific Conference in Taipei, Taiwan. In a subsequent workshop the merits were highlighted of using these methods in parallel: CEA, WTP, and ATP.
METHODS

The paper takes a 3-dimensional perspective (triangulation) in defining international pricing policy for pharmaceuticals using cost-effectiveness analysis (CEA), willingness to pay analysis (WTP), and ability to pay analysis (ATP). While each method has its own merits and can be used as stand-alone decision-support tool, the theoretical concepts on which they are based provide unique insight into efficiency, utility and equity. The objective is to find a balance between various economic methods of which some focus on effectiveness while others are geared towards incorporating equity in the equation.

Figure 1: International pricing policy framework and process
RESULTS

Method 1 – Cost Effectiveness Analysis (CEA)

Decisions regarding reimbursement and allocation of funds within the health-care budget are being influenced by the results of cost-effectiveness analysis (CEA) in an increasing number of countries. The term cost-effectiveness has become synonymous with health economic evaluation and has been used to depict the extent to which interventions measure up to what can be considered value for money.

Economic evaluation of health-care technologies essentially provides evidence on the efficiency of such technologies. Among other, such information might be of importance in several decision-making situations, including pricing and reimbursement. Different types of economic evaluations differ mainly in the way in which the effects of a health technology are measured. In cost-benefit analysis (CBA), costs and effects of a health intervention are both measured in monetary terms. Cost-effectiveness (CE) is often used in the field of health care, where it is difficult to monetize health effects.

The debate around measuring efficiency has shifted from using total or average cost-effectiveness ratios to incremental ratios (ICER) [1]. Each provides insight into the efficiency and the affordability of the intervention. The average CE ratio provides useful information about the overall affordability of an intervention. It is the net cost of a strategy divided by the total number of health outcomes. However, it is often more useful to examine the efficiency of one strategy relative to the other. This is done by calculating incremental CE ratios. To be able to compare interventions and capture their value, the numerator in the ICER must be expressed as a ‘single’ outcome, and therefore indicated in ‘natural’ units (for example, deaths avoided, or life years gained).

Nowadays, the preferential method for capturing health benefits is ‘Quality-Adjusted Life-Years’ (QALYs) which measures the health gained as a combination of the duration of life (years) and the health-related quality of life (QoL). The value assigned to quality of life is referred to as health utility and is further discussed in the next section. A combination of QALY outcomes is possible based on the total life years gained from a procedure and weighing each year to reflect the quality of life in that year [1]. The usage of QALYs as a standard measure for measuring health outcome in cost-effectiveness studies should strictly speaking be called cost-utility. Cost utility analysis (CUA) can be considered a special case of cost-effectiveness analysis, however, and the two terms are often used interchangeably.

When using QALY as the outcome measure, the ICER represents the ratio of incremental costs per QALY gained. The result of the calculated ICER outcome can be visualized on a cost-effectiveness plane consisting of four-quadrants [2] (Figure 2). The Y-axis usually captures the difference in intervention costs (less or more expensive), and the X-axis describes the difference in health effects. Outcomes positioned in Quadrant I (upper-right) are more effective and more expensive, those in Quadrant II (lower-right) are more effective and less expensive, those in Quadrant III (lower-left) are less effective and less expensive, and those in Quadrant IV (upper-left) are less effective and more expensive. In the latter case, when the new treatment is more expensive than the current treatment but does not lead to significant health gains, most policy makers may decide that this new treatment does represent value for money. Conversely, if the new technology is less costly but more effective
than the current comparator, the new intervention is described as dominant. The difficulty lies in assessing new technologies with an ICER that would position them in the upper-right quadrant where the new treatment is found cost-effective but the willingness-to-pay may reach a threshold beyond which payers are likely not to adopt the product. In other words, there is a threshold or ceiling ratio at which point the new product is considered no longer cost-effective. This threshold represented by the diagonal line indicates the maximum willingness to pay [3].

Figure 2: Cost-effectiveness and policy decisions

To summarize, cost-effectiveness ratios are an important criterion in health policy decisions. However, using CEA (and even CUA) alone may not capture all elements needed to make informed decisions on budget allocation do not tell whether there is a willingness to adopt the new health technology, drug or vaccine.

Method 2 — Willingness-To-Pay Analysis (WTP)

As stated above, to measure the willingness-to-pay (WTP) per QALY gained, the incremental cost-effectiveness ratio (ICER) must be compared with a 'threshold' or 'ceiling' value. Interventions with an ICER below this threshold may be considered for funding, whereas those with an ICER above the threshold tend not to be. Without explicating such a WTP threshold value, cost-effectiveness has limitations as decision-making tool, as it would lack a systematic and universally recognizable decision criterion.

Because of this standardized approach in measuring health outcomes, a comparison is subsequently possible between countries for example with regard to their willingness to pay per QALY outcome. Recent literature has seen a lively debate on implicit and explicit cost-effectiveness threshold(s), although without reaching consensus on the nature or height of an appropriate monetary value of a
QALY. At this moment, various institutions and governmental bodies have adopted threshold values in the process of optimizing the allocation of health-care resources, albeit sometimes implicitly and inconsistently.

How much money exactly is a payer prepared to pay for a gain of one QALY in the population? The acceptable ranges of the monetary value of a QALY used in such decision-making appear to be broad. Setting explicit thresholds may offer a range of theoretical advantages, however. It would for instance make the policy decision process more transparent and consistent. But health care decision-making is undeniably politically sensitive. Setting an explicit threshold almost certainly generates public debate about societal willingness-to-pay for health care. There is increasingly interest in the feasibility of modifying the ceiling threshold to reflect differences in the value society attaches to health gains depending on the characteristics of the individuals who receive them [4] [5].

Oncology is the therapeutic area in which the assessment of value and in particular willingness to pay is problematic. There is a public perception that cancer treatment is different, not necessarily from all other diseases, but from run-of-the-mill ones [6]. “That makes it harder to say ‘no’ to new treatments, which gets people into a discussion of how much clinical benefit new treatments need to demonstrate to justify reimbursement. This is a tough choice, however, because paying more for cancer care means less spending on other things. The rejection of given treatments because of WTP/QALY, opens up those who regulate formularies to accusations of passing death sentences.” [6].

In addition, decision-makers are not necessarily economists, and are reluctant to base their decisions on a single summary measure alone [7]. A ‘hard’ threshold approach dictates that results expressed for example as $/QALY, are taken prima facie and become the sole decision criterion. In contrast, adopting a ‘soft’ threshold makes room for considerations of other preferences. With this approach, the acceptability criterion does not lead to automatic acceptance or rejection but informs decision-makers [8]. Instead of a single figure, there is a threshold range with lower and upper boundaries. Interventions below the lower boundary will usually be accepted and made available; those above the upper level will usually be rejected, while the in-between outcomes fall in the zone of increasing discomfort as the cost goes “higher and higher” [8] [9]. This means that reimbursement of interventions falling between the lower and upper boundaries will be judged predominantly upon additional criteria.

Yet, the WTP/QALY measure is not entirely satisfactory from a theoretical and empirical viewpoint. First, the weighting of QALYs through either time-trade-off, standard gamble or visual analogue scale is highly subjective. Respondents usually do not pay for health care directly out of pocket but rather indirectly via insurance programs financed through taxes and/or employers, and thus their WTP values may be distorted by a lack of purchasing experience. Second, QALY outcomes focus exclusively on health outcomes, neglecting attributes of the treatment process such as fear, isolation, and confinement. Third, indirect societal benefits and costs are often not taken into account. Fourth, this measure does not allow pitting resources devoted to health against resources devoted to other uses. Specifically, it fails to reflect the preferences of citizens who may favour an expansion of the health budget, with the consequence that the threshold ICER value could be adjusted upward. This underlines the importance of further investigating the monetary value of a QALY.
The crucial question is whether QALYs can be defined meaningfully in a utility-theoretic model which represents lifetime preferences over health and wealth. Utility that is derived from the consumption of goods and services and thus indirectly from the monetary amounts spent was traditionally analysed in economic theory. This approach was called a welfarist approach. Another approach was termed ‘extra-welfarist’, and it was used to analyse non-good characteristics of individuals, e.g. the health status \[10\]. QALYs as a preference-based measure of sequences of health states are such a measure of extra-welfarism. We do not challenge the distinction between welfarist and non-welfarist as others have done \[11\]; but we share the view that one cannot separate individual preferences covering aspects of consumption from preferences covering aspects like health-related quality of life \[12\]. Preferences which cover both aspects are common in the literature, e.g. in the Grossman model which was developed to explain the demand for health care \[13\]. If preferences over health and consumption cannot be separated, QALY weights based on such preferences might depend on the consumption level. Then, different consumption levels might induce QALY weights that are biased in relation to each other. The utility of a health state has been shown to vary systematically on the basis of who is asked, how the question is asked, and how the health state is described.

An important issue to be addressed is about whose preference should be used in estimating the utility weights? The main contending sources of preferences are the general public (who are typically the ultimate payers), and patients. There are arguments for and against both groups, and some argue that the choice is a political rather than scientific one. There appears to be no theoretical basis for one unique WTP per QALY threshold \[14\]. However, there is consensus that despite these conceptual and pragmatic hurdles it would be good to continue research on defining the societal WTP/QALY. Another criticism to be dealt with relates to the fact that WTP/QALY does not fully take into consideration aspects of equity, and may therefore lead to biases related to an individual’s age and income level. When eliciting preferences amongst the young and poor, for example, WTP per QALY is likely to be very low. On the other hand, for the old and rich, WTP per QALY may be significantly higher \[14\]. It is hard to predict the overall effect of income from data in the literature, but the relationship is likely to be non-linear. A study that measured preferences in 3 patient populations using standard gamble, time trade-off, visual analog scale, and WTP to subsequently calculate WTP/QALY ratios, found that a natural log transformation of income was the best predictor of WTP/QALY \[5\].

In conclusion, the WTP/QALY is an important method for measuring value for money in health care. Nonetheless, the method fails to capture a number of factors that are potentially important and captures others with varying degrees of sensitivity. In spite of these hurdles, it is important to continue research in this area and work toward a higher level of transparency in societal decision-making.

Another method to guide international pricing decisions may be the analysis of the ability to pay (ATP) at the macro-economic level (nations) or at the micro-economic level for population-defined market segments. Especially if used in conjunction with differential pricing less affluent populations may be served at concessionary prices. The following section provides an overview on how both the gross domestic product and human development index provide additional insight and can be used as yardsticks.
Method 3 – Ability-To-Pay (ATP)

In this section, we argue that a country’s gross domestic product and living standards as they are captured in the Human Development Index (HDI) may function as a proxy measure of ‘real wealth’, and a basis for equitable, differential pricing between countries (whereas further intra-country differential pricing necessitates the supplementary use of house-hold income and health care system indexes).

One of the first proposals recommending using GDP as a yardstick in setting cost-effectiveness thresholds can be found in the report of the Commission on Macroeconomics and Health, which was commissioned by the World Health Organization (WHO) \[15\]. The report suggests that interventions costing less than three times GDP per capita for each disability-adjusted life year (DALY) (not necessarily QALY) averted would represent good value for money. Although this is still used as a rule of thumb, more sophisticated approaches have been developed to establish CE thresholds. The appeal of the GDP method is that it uses an objective national benchmark that is directly related to the affordability criterion. The concept of trying a national threshold to some type of economic objective benchmark is reportedly adaptable to QALYs \[8\]. As explained in section 2, methods based on WTP/QALY are not perfect either. Acknowledging the fact that different countries spent different amounts of their GDP on health care, would it not be more straightforward to measure instead of the ‘willingness-to-pay’ the ‘ability-to-pay’? And, would GDP be an appropriate benchmark to establish equity in pricing decisions?

We propose to use a combination of GDP-related macro-economic parameters with micro-economic value-based pricing methods. With regard to GDP, a more refined direct measure of a nation’s income is recommended, that is, the purchasing power parity (PPP) adjusted Gross National Income (GNI) per capita. An even more complete measure of wealth and economic development would be to use the human development index \[16\]. Both income-related indexes are based on large population surveys, made publicly available and updated annually by the World Bank and the United Nations Development Program, respectively. The Human Development Index (HDI) is a comparative yardstick that measures a nation’s social and economic development, and contains a log-based GDP formula. At the outset about 10 years ago, the HDI almost exclusively focused on the economic output of a country, but nowadays it applies a more holistic approach by including besides the standards of living, other important parameters like, for example, a population’s life expectancy, literacy rate and education. This in modern economics more accurately reflects a nation’s ‘real wealth’ and this is in the spirit of the ‘human capital theory’ which indicates that at times of advanced science and technology being the engines of economic progress, children and adults that are well educated and trained will determine economic progress, also in developing and emerging economies.

We propose to use the HDI curve as the basis for determining price discounts taking as anchor point a country with the high(est) GNI per capita at PPP reported, e.g. the US. This acts as value-based pricing benchmark from which to start calculating equitable prices in other countries (for which such evidence-based data are not always available). Alternatively, this benchmark value can be based on a basket of OECD countries where value-based pricing is being practiced. In this context, the HDI index serves as a price index that varies between 0% and 100%, whereas the market price takes into account the intrinsic value of the new or existing drug. The resulting equity curve is the graphic
representation of the concept of a price index calculated based on the country’s national income. Countries are ranked on the X-axis according to their national income, with the poorest countries to the left and the richest to the right. The ratio of wealth inversely reflects the theoretical discount rate that can be applied to each country compared to a high income reference country. The steepest discounts are reserved for the countries with the lowest income. Over time, as economies will grow, and convergence is taking place, especially in the BRIC countries, the discounts applied to low and middle-income countries are expected to decrease as the country’s wealth increases.

Figure 3: Pricing using HDI and GNI based indexes

Companies can compare the actual sales prices for existing products or predicted sales prices for new products against the ‘equity’ standard following the Human Development Index philosophy described above. The result can be calculated and displayed in two formats: a bubble chart and histogram. The bubble chart graphically shows a product’s price level (Y-axis) as well as the volume sold in each country during a particular year (bubble size), and plots these empirical data against the equitable price curve (blue). The non-linear regression connecting the sales input data gives an idea of how the empirical pricing curve (green) compares to the theoretical equitable curve, or deviates from it (figure 4).
In order to simply the method, the deviations can be highlighted in a bar chart highlighting prices that differ from the theoretical equitable curve (which depends on the margins tolerated) (figure 5). If the real sales price (bubble) matches the theoretical equitable curve, the ratio or index of real prices versus the theoretical standard for that country is 1 (which alternatively can be set at 100%).
To summarize, we recommend using the ATP method as an additional tool in international pricing decisions, and we have demonstrated that this makes sense from a societal fairness point of view. The concept of adjusting prices to buying power matches differential pricing theory which in turn is grounded in Ramsey pricing theory [17]. Whereas the term differential pricing stems from business economics, the term equitable pricing is derived from social welfare concepts. For more information on differential pricing, we refer to reference [18] which also describes models for intra-country pricing.
CONCLUSIONS

We have outlined a process for using three economic methods in conjunction (CEA, WTP, and ATP) in order to develop an evidence-based pricing and reimbursement strategy that is not only based on delivering proof of efficiency and cost-effectiveness but that also takes equity considerations into account.

The objective was to develop a strategy that provides the maximum possible access to new and existing medicines by covering virtually all citizens and patients across different economies regardless of their means while at the same time providing a return-on-investment to innovators and manufacturers to continue investing in developing new products that address unmet medical needs.

We recommend using all three methods. In other words, there is no one or two methods that are preferred while the third is not needed. Each viewpoint and associated methodology adds specific value and each analysis provides a deeper insight underpinning decision-making with regard to access and pricing. Although there is no clear hierarchy between the methods, the approach that we recommend is to build the triangulation process in a stepwise manner. In other words starting with an assessment of the intervention’s cost-effectiveness, followed by willingness to pay per QALY gained, and finally calculating the ability to pay by using a non-linear price index based on a country’s score in the human development index published by the United Nation’s Development Program.

This cascade of analysis in a sequential manner is made feasible by the fact that each method builds on the previous one. A calculation of the incremental cost effectiveness ratio is required before this outcome can be appraised against the ceiling ratio representing the maximum willingness to pay per QALY gained. This in turn defines the value-based price level at least in those countries where such methodology is applied. However, value based pricing requires availability of data about costs and benefits and not all countries have this. Although worthwhile this is no longer a prerequisite by adding the third method. If an anchor point is defined based on cost-effectiveness in a particular country where value-based pricing is practiced, or alternatively in a basket of OECD countries, the results in other countries around the world can be deducted by applying the equitable pricing method through the ATP model. Applying triangulation is therefore not about whether a particular method is superior and as a result one method should be preferred over the other, but by using these methods consecutively. At the end, this will provide a fuller picture than relying on one method.
ACKNOWLEDGMENTS

This study has been financially supported by Janssen, a Johnson & Johnson company. The authors were free in study design, collection, analysis, and interpretation of data, as well as writing and submitting the article for publication. The views expressed in this article are those of the authors and not the organisations they work for.
REFERENCES


