Volume 9 • Number 3 • 2006
VALUE IN HEALTH

# What Decision-Makers Want and What They Have Been Getting

Maurice McGregor, MD

McGill University Health Center, Royal Victoria Hospital, Montréal, QC, Canada

## Introduction

I believe we would all agree that the only reason we produce health technology assessments (HTAs) is to inform health policy decisions, and when they fail to do so they are a waste of time and effort. Nevertheless, at the present time it seems that HTAs often have little impact [1–4]. My first objective will be to consider some of the factors that cause HTAs to fail to influence policy. My second will be to report on an ongoing experiment in HTA development that attempts to minimize these factors, and has resulted in an extremely high impact of HTA reports on health policy decisions.

Before I start, I have two caveats. First, my remarks are *not* evidence-based. They are reflections resulting from my reading and my personal experience both as an HTA producer and as a health-care decision-maker. Second, my experience is mostly Canadian, so some of my reflections may not be completely relevant to other countries.

Furthermore, the decision-makers I will be referring to are the meso-level policymakers, the people who make technology acquisition decisions at the hospital or regional level. In our Canadian system of universal health insurance, top-level policy, such as decisions on what programs will be funded, are taken at government level. Increasing the impact of HTAs on such decisions has been discussed elsewhere [5]. But hospitals or regions mostly take meso-decisions, such as those on the acquisition and use of new technologies. These organizations are usually given a capped budget and expected to conform to the Canada Health Act, according to which they must to give all "necessary" treatments to all Canadians, a feat that clearly requires careful prioritization. Sound prioritization decisions require well-prepared, objective evidence such as is found in HTAs. So if HTAs are not influencing these decisions it is urgent that we consider what to do about it. Our options are limited.

I believe the UK solution to this problem is not an option for us in Canada. In the UK the final path of all HTAs is the National Institute of Clinical Excellence (NICE), and their policy "recommendations" are now virtually mandatory, and Regional Authorities are required to follow them. North America is probably too big and regional differences too great for such a solution to be accepted. So we must consider the other option. We must ask why many HTAs lack impact and what we can do about it.

10.1111/j.1524-4733.2006.00098.x

## Why Some HTAs May Lack Impact

For HTAs to have impact, they must first of all be understood and be acceptable to those who use them. But decision-makers do not always find this easy. They have problems, both with their content and with their presentation.

### *Content Issues*

The first problem is acceptance of the quality-adjusted life-year or QALY [6]. This is at first surprising. The concept is immensely attractive. When it is your job in a hospital or regional authority to have to allocate resources with a capped budget, in theory it is a wonderful solution to be able to compare quite different health outcomes—curing heart disease or headaches—using a single unit such as the QALY. But many decision makers do not trust this unit. They have questions, such as these:

*Is it meaningful?* It may well be meaningful to health economists, but the institutional administrators who have to use them have difficulties. Can interventions that reflect the quality of life really be measured in units of length of life? Can we really estimate the health benefits of Vioxx or Viagra, which have no effect on longevity in units of length of life survived?

The following example taken from a study of Smith and Roberts [7] helps to illustrate their dilemma. This report concluded that the cost-effectiveness of taking five Viagra pills per month compared with no therapy was $11,230 per QALY. Because this was less than the cost of a year of life saved by cholesterol lowering medication, or coronary artery bypass surgery, and in addition was below the critical threshold of $50,000 per QALY, Viagra should be considered a higher priority in the allocation of resources than treatment of these medical conditions.

Now many decision-makers would say any analysis that arrives at this sort of conclusion needs to be ques-

 

tioned. And of course questions are not lacking. They relate in particular to the measurement of the health status preference, or utility, which when multiplied by life expectancy constitutes the QALY. These questions are not new. They have been extensively debated before. Yet, QALYs are still used to inform policy decisions as if such questions did not exist, or had been successfully answered. For example:

*Whose utility?* One of the oldest questions is *whose* health preference (utility), should be measured when making decisions that involve shared resources? Smith and Roberts did not actually measure any. They used the findings of Volk et al. [8] who, a year or two earlier had carried out a study of 10 healthy middle-aged men. These men, when asked through time trade-off what proportion of their remaining lives in perfect health they would give up to avoid impotence, came up with an average response of 0.26. Health professionals have arrived at lower estimates such as 0.05 [9] and 0.15 [10], while the wives of the 10 men studied by Volk and colleagues arrived at a value of 0.02 [8]. So the issue is this. Should we be measuring utility as estimated by the sufferers, by their wives, by average citizens, by taxpayers, or by health professionals? The answer is, we do not yet know. But until we all adopt the same approach it is clear that for this reason alone, this unit, and the QALY derived from it, is an inappropriate measure to use as a determinant of serious health policy decisions.

*Which utility?* Decision-makers are increasingly aware that utilities are measured by different techniques such as time trade-off, and standard gamble. Nevertheless, these techniques do not consistently give the same results [11]. Which technique should we believe? Again, we do not know [12]. But clearly for serious decision making we should not be comparing QALYs derived by one technique with QALYs derived by another.

*Reliability?* Unfortunately, the measurement of utility is not consistently repeatable; when it is carried out by different investigators on the same subject, by the same investigators on different subjects [13], and even by the same investigators on the same subjects [14]. Decisions based on an unreliable measurement should surely not determine serious health policy decisions.

*Threshold?* It is commonly presumed, as in the Smith study quoted above, that there is some threshold such as $20,000, or $30,000, or $50,000 per QALY that separates acceptable from unacceptable technology acquisitions. Useful as this would be, there is unfortunately no logical reason to support any particular generalizable threshold. (Any health authority can of course arbitrarily decide that "x" is the upper limit of what *they* are willing to pay, but their decision has no validity outside the jurisdiction in which it is made).

*The role of QALYs in decision-making.* The sole objective of cost utility analysis is to enable comparison of the price of achieving different health benefits, either explicitly in a league table or implicitly, when trying to establish priorities. If the unit of comparison varies according to who estimates it, and the way in which it is estimated, and if it is in addition unreliable, then it should clearly not be used as the a determining factor when making policy decisions on the acquisition or use of health technologies.

Fortunately, it turns out that in the real world of meso-decision-making, cost-effectiveness seldom determines policy decisions. In practice the first and principal determinant is the estimated health impact of the technology in question, and the second is the opportunity cost (what will have to be given up to pay for it). Nevertheless, the opportunity cost is often difficult to identify. When an institution spends $1 million on coated stents or implantable cardiac defibrillators the money is usually taken off the top of the institutional budget, and the budgets of all departments in the institution are squeezed proportionately. So as a surrogate for opportunity cost, decision makers use the estimated budget impact of a technology acquisition (expenditure by the institution as a fraction of institutional budget).

These two factors, the health impact and the budget impact are issues of objective scientific inquiry. By contrast the third determinant, desirability, is very different. Whether the available resources should be spent on relief of headaches, hemodialysis, or heart surgery, is a subjective decision, dependent on political and social pressures and the opinions and values of the decision-makers. Accordingly, such decisions are not generalizable. They are not the same in Beverly Hills or Bangladesh and because they are locally variable it is helpful to develop local ways of arriving at these decisions.

In practice cost-effectiveness provides the decision makers with a sense of the relative value of the acquisition in question. It tells them whether or not they have a "good buy." The knowledge that a technology that they want and need is available at a low cost in

Determinants of hospital level decisions
1. **Health impact:** The size of expected health benefit resulting from expected use.
2. **Opportunity costs:** What they must do without. Because hard to identify, we use as a surrogate.
**Budget impact:** Net Cost/Institutional budget.
3. **Desirability:** How much it is wanted/needed.
4. **Cost-effectiveness:** Is it good value for money?