Health economics and health technology assessment

Article · May 2018

DOI: 10.1016/j.mpmed.2018.04.002

CITATION
1

READS
77

1 author:

Anthony John Culyer
The University of York
399 PUBLICATIONS **7,130** CITATIONS

SEE PROFILE

Health economics and health technology assessment

Anthony J Culyer

Abstract

Health technology assessment (HTA) is a set of techniques for helping decision-makers make better decisions in terms of the impact on a nation's health. This article outlines the principal features of HTA; it emphasizes the basic economic analysis that lies behind it, the integration of economic and other evidence, the ways in which it is best conducted, and its role as an aid to, rather than a substitute for, careful thinking.

Keywords Cost-effectiveness; cost-effectiveness threshold; efficiency; fairness; health technology assessment (HTA); NICE; opportunity cost; quality-adjusted life-year (QALY)

Key points

- Health technology assessment (HTA) is, for the UK, a way of working out the ways in which NHS resources will have the maximum impact on the nation's health
- It helps decision-makers to balance considerations of efficiency and fairness
- It should never be seen as a short cut, or a way of avoiding difficult decisions, in multidisciplinary and multiprofessional contexts
- It should always be conducted in a participative and deliberative fashion
- All stakeholders ought to understand the basic principles underlying HTA so they can commission good work, detect incompetence and give their approval when all is as it should be

Introduction

Health technology assessment (HTA) is a set of techniques that helps decision-makers make better decisions. In this context, 'better' means that the resultant decision is likely to be more in accordance with two things: (1) the decision-maker's objectives; and (2) a reasonable understanding of the evidence.

A fundamental assumption of HTA in the UK is that at least one of the decision-maker's objectives is to use National Health Service (NHS) resources so they will have maximum impact on the health of the population. However, because improving patients' health is rarely an exclusive objective, it is often necessary for HTA to consider other consequences of decisions. These include the financial impact on individuals and their families, and the fairness of how the decisions distribute health and lifetime well-being among the NHS's clients. Several immediate implications flow from these basic elements of HTA, as described below.

Identity of the decision-maker

The decision-maker's identity must be clear. Are they, for example, a government minister, the chief executive officer (CEO) of an insurance agency, or someone accountable to these senior members of an organization? The decision-maker is the ultimate authority on matters relating to the objectives and to questions of ethical value, all of which must be practically and politically feasible.

In the UK, the ultimate policy decision-maker is the Secretary of State for Health together with public agencies that have decision-making powers, for example the National Institute for Health and Care Excellence (NICE). Non-public agencies, like employer organizations, manufacturers and trade unions, can also commission HTA studies to inform their own decisions.

Organization of thinking

HTA is a way for decision-makers to organize their thinking about the choices they have to make. HTA is not a substitute for thought but, if done well, will provide a practical framework for that thinking. This can include minimizing the risk of double-counting or making significant omissions, maintaining a logical coherence and clarifying the kind of evidence needed, which might be qualitative or quantitative.

Multidisciplinary nature

HTA is necessarily multidisciplinary. The main disciplines involved are clinical medicine (depending on the context), epidemiology, economics, biostatistics, and systematic reviewing and meta-analyses.

Logical framework

HTA provides a consistent logical framework embodying two contrasting elements. The first is *science-based*. This uses reliable theory-based tests to compare hypotheses against empirical evidence, most of which is medical and economic in nature. Empirical evidence, however, can also be behavioural, for example when the outcomes of decisions are contingent on the behaviour of clinicians, managers or patients.

The second element is *value-laden* and requires social value judgements to be made. These include what is understood by 'perspective', 'health' and 'fairness', and how to weigh things up when objectives are in conflict or cannot be simultaneously realized.

Anthony J Culyer *CBE BA Hon DEcon Hon FRCP FRSA FMedSci is an Emeritus Professor of Economics at the University of York and Visiting Professor at Imperial College London, UK. Competing interests: I have acted as an expert witness on health economics for an author writing a history of the subject.*

The basic analysis

HTA is basically a simplified version of 'reality' — it is not intended to be a complete or realistic description, but it is meant to contain features that allow important distinctions to be made, and suggest the type of evidence that might be gathered.

Outcomes

Central to HTA is the analysis of alternatives: HTA always involves a comparison between alternative interventions. If decision-makers cannot make reasonable comparisons, they can hardly make reasonable choices. This may seem self-evident, but a very large number of effectiveness studies quote measures of outcome that are disease-specific, meaning that comparisons with other disease outcomes are then difficult.

Because a central objective of healthcare is to have an impact on health, a generic measure of this impact is required that should work reasonably well for many different types of intervention. One measure used by NICE is the quality-adjusted life-year (QALY). This defines health in terms of a person's physical and mental attributes.

QALYs

Constructing a QALY is straightforward in principle. Sampling the population yields weights for various 'dimensions' of health. The frequently used EQ-5D-3L (EuroQol Group — 5 Dimensions — 3 Levels) version has five dimensions: Mobility, Self-care, Usual activity, Pain/discomfort, and Anxiety/Depression. Each of these can be reported at three levels: no problem, some problem, extreme problem. The EuroQol Group (https://euroqol.org/) has also developed a five-level version (EQ-5D-5L), as well as an instrument validated for self-completion by children and young people (EQ-5D-Y). A life-year of full health gets a QALY score of 1, while being dead gets a QALY score of 0. Obviously, comparing outcomes in terms of QALYs will not necessarily cover all relevant aspects of outcome in all cases. It is therefore advisable for the decision-making process to have a 'reality-check' so that patients and their families can draw attention to any possible omissions or other biases in the data.

Efficiency

An efficient allocation of the NHS budget exists when it is impossible to increase health outcomes for some patients without reducing outcomes for others. Figure 1 illustrates how choices about the services to be provided (or not) ought to be considered. A major simplification in this basic model is that it does not take account of uncertainty, for example how adequate the outcome measure is, or how effective the various interventions are. For convenience, it is also assumed that there are a rather small number of interventions and a single time period. These are all complications that more advanced analysis tackles.

Imagine a bookshelf of healthcare interventions; each intervention is a book, and each is ranked by height according to its effectiveness. In this example, 'effectiveness' means the health gain in QALYs generated per £1000 of NHS spending. The most effective intervention — tallest book — is positioned on the left, and the less effective ones stretch away on the right.

The fatness of each book represents the estimated cost of providing the intervention. This fatness is a combination of several things: (1) the cost of a specific technology, such as a drug; (2) the costs of associated procedures (other medicines, diagnostic services, community services, etc.) for as long as the treatment continues; and (3) the estimated number of people using the intervention. The area of each book's spine thus measures the anticipated expenditure on each intervention.

To maximize the impact of the NHS budget on health, the decision-maker ought to select the first book on the left and then add books (i.e. further interventions) along the shelf until the money runs out. At that point, all the interventions included will be effective, and only the most effective of the effective ones will have been selected. The only services to be offered by the NHS will therefore be those to the left of the 'budget limit' line. The least cost-effective intervention that is included defines the effectiveness-cost 'threshold'. If turned upside down, this effectiveness-cost threshold becomes the cost-effectiveness threshold (currently around £20,000 as used by NICE). An increase in the NHS budget enables more technologies to be used, or existing ones extended to more patients, and the budget limit moves to the right. This means in turn that the NICE cost-effectiveness ratio can rise with the NHS now providing services that were previously judged insufficiently cost-effective.

So why are all effective interventions not to be provided? It is not because they are ineffective — on the contrary, all the interventions on our bookshelf are effective; one would have to go a long way to the right before hitting zero productivity. The trouble with the ones not being used is that they are *not effective enough*. The benchmark test for including further interventions is the cost-effectiveness of the *least* cost-effective intervention that is included. This cost-effectiveness is the impact per £1000 that has to beaten.

What needs always to be demonstrated in trials and other assessments of clinical productivity is relative (rather than absolute) effectiveness. One way of doing this is to make direct comparisons between interventions, such as comparing alternative treatments for cancer, or for macular degeneration. A less cumbersome procedure is to use the cost-effectiveness threshold and make comparisons with that.

Another implication of the model is that the threshold and the budget are intimately linked. This is because what determines the threshold are: (1) the productivity of the interventions (i.e. their impact on health), and (2) the size of the budget. If the CEO of NHS England were to complain that NICE approves interventions that cannot be afforded, the model tells us that either:

- the threshold used by NICE is too generous
- the budget for the NHS is too small
- government policy is inconsistent — giving the impression of wanting to spend more but not providing the necessary funding.

Opportunity cost

It is common these days to hear references to 'opportunity cost' in healthcare finance and budgeting. Opportunity cost is an important idea. Adding additional interventions to the available NHS financial package obviously must mean some displacement of other interventions. In a perfect system, the monetary value of the threshold would be the same as that of the least productive intervention offered in the NHS. Any innovation with a lower cost per QALY than the NICE threshold should then replace that intervention, and the NHS budget would thereby generate more.

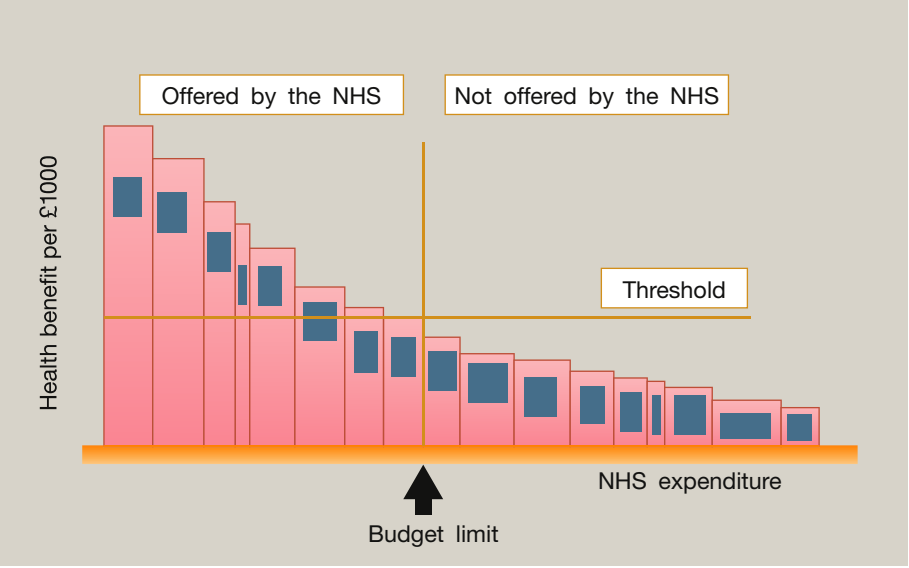

Figure 1 The Bookshelf of technologies.

The health no longer generated by the displaced intervention is the opportunity cost of introducing the innovation, which should always be lower than the health gain from the new treatment – otherwise there will be a negative impact on population health. Opportunity cost in healthcare decisions is thus the heath forgone by choosing one technology rather than another, or using a technology on one subgroup of patients rather than another. Opportunity cost focuses the mind on the fact that cost is not simply money, rather it is the health that could have been bought with that money. Can this cost be justified? The logic of HTA says yes, but only if the resources in question generate more health than the health forgone as resources move from one use to another.

Suppose that the health (QALYs) lost through squeezing mental health services in order to accommodate new cancer treatments was larger than the health gained from the cancer treatments. In this situation, either the threshold is failing to indicate the true opportunity cost of the new treatment, for example by weighting the gains to cancer patients higher than those needing mental health services, or decision-makers are ignoring it and thereby ensuring that population health (measured in QALYs) is lower than it can be.

In the end, the decision remains a matter of judgement – sometimes a judgement about the facts, and sometimes a judgement about what is right and fair. But being right and fair is as much about the rightness and fairness of opportunity costs as it is about the rightness and fairness of the direct health benefits of treatments.

Fairness

NICE's usual assumptions involve a kind of social equality between QALYs: that each QALY counts for the same regardless of who gains it, how many they gain, their existing health and their past history. This assumption is, however, a social value judgement rather than a science-based requirement. The only science that may be involved arises when surveys of public opinion on the matter are conducted. HTA analysts consequently need to establish what weights are to be embodied in overall community outcome measurements. Is a QALY gain for a chronically sick person near death to have the same weight as a QALY gain for an otherwise healthy youngster with normal life expectancy? This social value judgement is not for HTA experts to decide; it is for policy decision-makers. HTA experts can help by pointing out the possibilities and their logical consequences. Technical skills alone, however, do not entitle the experts to make decisions on behalf of society – unless, that is, they have been appointed to do just that and as such are publicly accountable for their decisions.

Giving a greater weight to some patient groups may seem a simple matter, but it can have complex consequences. For example, when evaluating the proposed introduction of a new late-stage cancer medicine, decision-makers may take the view that people near the ends of their lives should receive a greater weight. This makes it more likely that such a new product will pass the threshold test. However, it is highly unlikely that the cancer patients in question are the only patients near the ends of their lives. Many who are already receiving NHS care may also be near death, and their QALYs ought also, on grounds of horizontal justice (in other words, treating people who are alike in like ways), to have a higher weight. If this affects existing interventions, this will increase the effectiveness-cost threshold (cause the cost-effectiveness threshold to fall). However, it may also warrant increases in the resources going to other non-marginal interventions, with the consequence that, conversely, the cost-effectiveness ratio will rise! A well-informed judgement is required – the matter is not as simple as may have been initially thought.

In the UK, because treatments are essentially free of charges to the patient, there are very few complicating issues associated with evaluating the efficiency or fairness of interventions for which the patient must fully or partly pay. However, decision-makers might want to favour interventions that could have an equalizing impact on the distribution of the burden of disease

according to, say, geography or socioeconomic status. In that case, HTA would need an expanded evidence base so that estimates of the impact of a new intervention on the relevant distribution could be made; this would then be compared with alternative possible new interventions and the distributional impact of the interventions that have been displaced.

Taking equity factors into account does not mean that there is a trade-off between efficiency and fairness. Whatever the weights assigned to QALYs, an efficient resource allocation still exists when no person's weighted QALY can be increased without reducing someone else's. It follows that there are in principle as many fully efficient allocations of the NHS budget as there are possible QALY weights. So, being efficient in the NHS does not require all QALYs to be treated the same, even though that is a social value judgment that commands wiespread support. What *is* a condition of efficiency is that the weight(s) assigned to QALYs should be explicit and judged to be socially appropriate. That done, the twin goals of efficiency and fairness can be delivered together!

Other considerations: process and technical

Evidence of any kind is rarely either complete or completely relevant. Sometimes it is altogether lacking. Often it is disputed. The quality of the evidence therefore needs to be judged. Value judgements are also inherent in HTA. Decisions about the interventions to be provided by the NHS will affect some interested parties for better, but others for worse. For all these reasons, the process of conducting an HTA should be open, as transparent and as participative as possible.

Because HTA is a way or organizing thinking, much of the process is necessarily deliberative; stakeholder groups can be represented in this and provide checks. These checks include:

- whether the right comparisons between products and interventions are being drawn
- whether the measures used (e.g. QALYs) are corresponding acceptably to the practical experience of patients and their families
- whether the scope of both costs and consequences is being outlined appropriately (are costs falling outside the NHS to be included?)
- whether issues of feasibility are being addressed
- whether the evidence is being professionally reviewed and appraised
- whether one or more stages in the process may have been affected by someone's vested interests.

An early task in any HTA is therefore to make a list of the relevant stakeholders and decide when and how their representatives might most appropriately be involved.

Final comment

It is not necessary for most stakeholders, including clinicians, to grasp fully the technical aspects of HTA. They should, however, understand enough about the method of analysis, and the processes of actually conducting it, to be able to ask searching questions, detect incompetence or conclude that all has been done as well as could be reasonably expected. ◆

FURTHER READING

Claxton K, Paulden M, Gravelle H, et al. Discounting and decision making in the economic evaluation of health-care technologies. *Health Econ* 2011; **20:** 2–15.

Claxton K, Sculpher M, Palmer S, et al. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? *Health Econ* 2015; **24:** 1–7.

Culyer AJ. Cost-effectiveness thresholds in health care: a bookshelf guide to their meaning and use. *Health Econ Pol Law* 2016; **11:** 415–32.

Paulden M, O'Mahony JF, Culyer AJ, et al. Some inconsistencies in NICE's consideration of social values. *Pharmacoeconomics* 2014; **32:** 1043–53.

Wilkinson T, Claxton K, Sculpher MJ, et al. The International Decision Support Initiative reference case for economic evaluation: an aid to thought. *Value Health* 2016; **19:** 921–8.