Clinical effectiveness assessment in HTA

Introduction

Assessing the impact of any technology requires comprehensive information that reflects what is likely to happen in a health system or society. Good analysis requires the use of expert advice and methods from the various disciplines that are used as inputs.

HTA assessments are comparative analyses, which compare the existing standard of care with the new technology to see what value the new technology would provide (the so-called 'added value'). Countries, regions, and hospitals implement HTA assessments in different ways. All will consider the health problem in their local context and then assess the treatment for the indication that the regulators have agreed for the medicine.

Within that indication, HTA bodies examine available data to evaluate how well the treatment can work in comparison with the best standard of care (in terms of safety and clinical effectiveness). Some HTA bodies also assess the costs and cost-effectiveness of a medicine, and while some do formal assessments of the ethical, organisational, social, and legal aspects, others simply consider these issues implicitly in their assessment.

The 'added value' of a technology is determined by each HTA organisation in a different, multifaceted way. Conclusions about the 'added value' of a technology may differ between HTA organisations. The European network for HTA (EUnetHTA) has developed a framework by which the 'added value' can be assessed, called the HTA Core Model®.[1] There are nine domains

in the HTA Core Model®:

1. Health problem
2. Technical description of technology
3. Safety
4. Clinical effectiveness
5. Costs and economic evaluation (cost effectiveness)
6. Ethical analysis
7. Organisational aspects
8. Social aspects
9. Legal aspects

EUnetHTA has defined the assessment of the first four domains as a 'relative effectiveness assessment' — in which, for instance, a new treatment is compared to the existing treatment(s).

Clinical effectiveness assessment in HTA

A clinical effectiveness assessment investigates the effect of a new technology on the health of patients in a standard clinical setting compared to that of the current standard of care. The impact that a technology has on health is usually analysed through a further examination of health outcomes. Patients want access to new medicines that:

- reduce outcomes that are perceived as 'bad' — such as heart attacks, hospitalisations, and side effects, **and/or**
- increase outcomes perceived as 'good' — such as improved functionality and pain-free days.

During a clinical effectiveness assessment, HTA bodies use established methods from associated disciplines of medicines. In particular, the evaluation of clinical effectiveness is carried out with principles borrowed from epidemiology and medicine (called 'clinical epidemiology').

There are four underlying principles of good clinical

effectiveness assessment:

1. Seeking information,
2. Asking relevant questions,
3. Understanding differences, **and**
4. Valuing differences.

Seeking information

HTA bodies use clinical information to estimate what health outcomes patients might experience when given a new medicine. First, though, they must decide how they will gather the information. There are three different ways HTA bodies might obtain clinical information on new technologies:

1. Reviewing existing information on the medicine's performance,
2. Conducting a new study to gather information and evaluate the performance of the medicine in a real-world setting, **or**
3. Asking clinicians and patients ('experts') what their expectations are of the medicine.

HTA bodies often use a combination of these approaches. For example:

- They might use information from the technology's Marketing Authorisation Holder (MAH) to inform their own independent reviews and analyses.
- When information is missing, expert opinion may be required — for example, to find out whether changes in short-term outcomes (such as lowering cholesterol) might predict changes in longer-term outcomes (such as avoiding hospitalisation).

HTA bodies rarely commission new studies are rarely commissioned, because the time required to set up and have a study approved is typically too long. In some cases, the responsible bodies have allowed a medicine to be reimbursed

conditionally, based on the collection of further information (this would be similar to regulatory authorities granting a conditional Marketing Authorisation (MA) requiring further information to be gathered). The risk of a new medicine performing poorer than expected in the real world can then be shared between the MAH and responsible body through price negotiation mechanisms or other changes to the conditions associated with reimbursed access (such as further restrictions on the patient population eligible to receive reimbursed access) while patients are given more immediate access.

Asking relevant questions

When assessing the clinical effectiveness of a new health technology, the HTA body must be careful to consider all outcomes associated with it. It is important to know about these outcomes in order to ask the relevant questions about the technology's effectiveness.

There is increasing understanding that the outcomes that may seem important to clinicians are not always those considered most important by patients. For this reason, it is important for patients to be involved in designing studies, in order to make sure that information is collected on the outcomes that matter to them. For instance, in recent years it has been recognised that quality of life is an important outcome for patients. This has led to the development of specific methodologies to create quality-of-life measures and so-called 'patient-reported outcomes' within clinical studies.
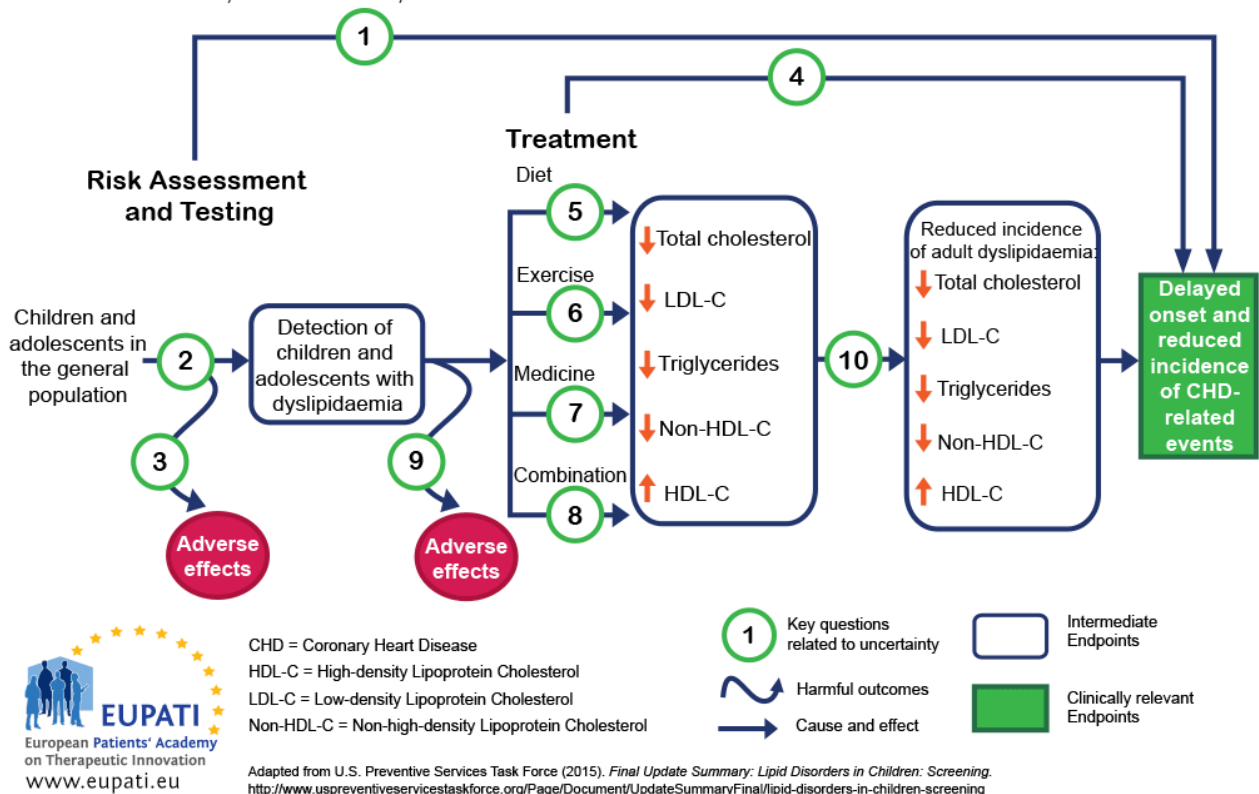
One approach to ensuring that all the important outcomes of a particular technology are examined is to use an analytic framework — for instance, a flowchart as in Figure 1.[2] Analytic frameworks are helpful to visualise all of the outcomes associated with an intervention, and to highlight where there are uncertainties.

In the analytic framework in Figure 1:

- Cause and effect are depicted by arrows.
    - Curved arrows indicate harmful outcomes.
- Health improvement outcomes (such as decreased mortality) are depicted by rectangles.
    - Sharp cornered rectangles show clinically relevant endpoints (those that are perceived by the patient, such as chest pain)
    - Round cornered rectangles show intermediate endpoints, including surrogate endpoints (which cannot be perceived by the patient, such as cholesterol level in blood).
- Key questions related to uncertainty can then be shown numerically.

Example of an analytic framework

Screening and treatment of children and adolescents for dyslipidaemia: Interventions, outcomes, and adverse effects

CHD = Coronary Heart Disease
HDL-C = High-density Lipoprotein Cholesterol
LDL-C = Low-density Lipoprotein Cholesterol
Non-HDL-C = Non-high-density Lipoprotein Cholesterol

Key questions related to uncertainty
Harmful outcomes
Cause and effect
Intermediate Endpoints
Clinically relevant Endpoints

EUPATI
European Patients' Academy on Therapeutic Innovation
www.eupati.eu

Adapted from U.S. Preventive Services Task Force (2015). *Final Update Summary: Lipid Disorders in Children: Screening.* http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/lipid-disorders-in-children-screening

This analytical framework was used to determine the strengths and limitations of evidence for the effectiveness of screening children and adolescents for

dyslipidaemia (disorders of lipid metabolism) as a part of routine primary care. Dyslipidaemias are important risk factors for coronary heart disease (CHD).

The key questions related to this analytic framework are the following:

- **Key Question 1:** Is screening for dyslipidaemia in children/adolescents effective in delaying the onset and reducing the incidence of CHD (coronary heart disease) related events?
- **Key Question 2:** What is the accuracy of screening for dyslipidaemia in identifying children/adolescents at increased risk of CHD related events and other outcomes?
- **Key Question 3:** What are the adverse effects of screening (including false positives, false negatives, labelling)?
- **Key Question 4**: In children and adolescents, what is the effectiveness of drug, diet, exercise, and combination therapy in reducing the incidence of adult dyslipidaemia, and delaying the onset and reducing the incidence of CHD-related events and other outcomes (including optimal age for initiation of treatment)?
- **Key Questions 5-8:** What is the effectiveness of drug, diet, exercise, and combination therapy for treating dyslipidaemia in children/adolescents (including the incremental benefit of treating dyslipidaemia in childhood)?
- **Key Question 9:** What are the adverse effects of medicine, diet, exercise, and combination therapy in children/adolescents?
- **Key Question 10:** Does improving dyslipidaemia in childhood reduce the risk of dyslipidaemia in adulthood?
- **Key Question 11** (not pictured): What are the cost issues involved in screening for dyslipidaemia in asymptomatic children?

Understanding the differences between outcomes

Once all important outcomes are identified, there may still be several challenges in comparing the effects of a new technology with the standard of care and other existing treatments. Outcomes may be measured in different ways, or two technologies may seem to have similar outcomes until closer inspection shows differences.

In cases where the important identified outcomes are difficult to measure, or have never been measured before, scientists must carefully create a measure that can then be reproduced in a study. For instance, a patient may want to know how a medicine will help them return to work or get out of bed. Scientists may create a numeric pain-rating scale for patients with lower back pain. In other cases, for instance where a study measures a change in a laboratory parameter, this change needs to be re-interpreted into a measure that matters more to patients — such as the ability to return to work.

Sometimes, regulators who approve medicines may be satisfied with a medicine's manufacturer demonstrating the effect of a new medicine via a short-term outcome, such as lowering blood pressure. An HTA body will need to re-interpret that short-term outcome into more patient-relevant outcomes, such as avoiding premature death.

Some outcomes may seem intuitive but, upon closer examination, may be difficult to interpret. For example, a reduction of the risk of five-year mortality (death within five years) by 50% does not mean that the medicine can prevent premature death. It could simply be:

- extending life expectancy from 4.9 to 5.1 years (or worse, 4.99 years to 5.01 years) in some patients, **or**
- curing disease in very few but not extending life at all in others.

Even if differences in measures that are meaningful to patients are observed, these may still be difficult to interpret. For example, studies may indicate a new medicine reduced the risk of hospitalisation from infection by 33%. However, this may mean different things. It could mean that:

- 33 out of 100 people taking the medicine who would have otherwise been hospitalised avoided hospitalisation (this is called an **absolute** reduction in risk), **or**
- the chance of hospitalisation is reduced by 33% relative to the chance of being hospitalised without medicine (this is called a **relative** reduction in risk). If the chance of being hospitalised in the absence of the new medicine is 3 out of 1,000, then a 33% reduction reduces this to 2 out of 1,000. This means 1 out of every 1,000 people taking the medicine will benefit. This is quite different from the 33 out of 100 people benefitting in the example above.

A final challenge to understanding the differences between a new health technology and the standard of care is the use and misuse of statistical tests. Statistical tests are intended to help researchers know if the differences they have detected are likely to be real. Often, this is reported in the form of a p-value. However, p-values do not reflect the magnitude (size) of the difference, or whether that difference is meaningful to patients. This means that p-values are generally not useful for patients and providers making decisions.

Other statistical measures include confidence intervals. Confidence intervals are more helpful, because they give some sense of the size of the difference between the new health technology and the standard of care. Confidence intervals also reflect any uncertainty about the estimate of the magnitude of difference. For example, a new medicine may be reported to reduce the chance of having a future heart attack by 33% (with a 95% confidence interval of 5% to 45%) relative to the current chance of having a heart attack.

Valuing differences

The last challenge is to understand how to perceive and value the differences between outcomes. If a medicine prolongs life by 0.2 years, HTA bodies still need to know:

- how much a patient would value 0.2 years of additional life given expected side effects and other concerns,
- if all patients experience roughly the same gains or if there are dramatic differences in patients, **and**
- if all patients value these gains similarly.

A new medicine that increased life expectancy by an average of 0.2 years would be perceived differently if it worked in some patients but had no effect on others, when compared with a scenario where all patients gained 0.2 years with little differences across patients.

There are several mechanisms that can be used to understand the relative value that patients and providers put on differences in health outcomes. One is qualitative research, such as surveys or focus groups, intended to provide an understanding of which outcomes are most important to patients. Another is quantitative research based on surveys of patients, which can assign precise numerical values to the importance placed on different states of health.

In short, an assessment of clinical effectiveness should address the following questions:

1. How comprehensive was the information?
2. How accurate is the information?
3. Is anything missing?
4. How understandable is the information?

References

1. *HTA Core Model*. Retrieved 7 December, 2015, from

http://www.eunethta.eu/hta-core-model

2. U.S. Preventive Services Task Force (2015). Final Update Summary: Lipid Disorders in Children: Screening. Retrieved 7 December, 2015, from: http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/lipid-disorders-in-children-screening

A2-6.03.1-v1.1