TNO Prevention and Health

Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek/Netherlands Organisation for Applied Scientific Research

B 20

Division Public Health
Wassenaarseweg 56
P.O. Box 2215
2301 CE Leiden
The Netherlands

www.tno.nl

T +31 71 518 18 18
F +31 71 518 19 20
Info-VGZ@pg.tno.nl

TNO report

PG/VGZ 2002.034

Working Paper 6

Creation of the health benefits package proposal with specification of groups of services with the best and worst efficiency parameters

Date	February 2002
Authors	H.D. Banta K. Landa
Project name	Standard of health services purchased in the national health insurance system (contract nr. 3.4.1.41) - Poland
Project number	011.41053

Introduction

The Polish Ministry of Health and Social Welfare states that "there have been a number of misconceptions in the area of responsibility of the payer, service providers or territorial self-governments for the health care . . . These misconceptions directly result from lack of determination of the so-called health benefits package which the payer . . . is obliged to purchase for individuals under his care (the insured)" (Terms of Reference, 1999).

The Netherlands Organisation for Applied Scientific Research (TNO), in its contract with the Polish Ministry of Health, stated that the Ministry of Health has acknowledged a number of misconceptions that directly result from "lack of determination of the so-called health benefits package which the payer in the form of a public insurance institution is obliged to purchase for individuals under his care (the insured)." This lack of determination of the benefits package is a principle barrier to increasing the efficiency and effectiveness of health care delivery in Poland.

The purpose of the TNO project is to assist the Polish Ministry of Health with this problem. The TNO project proposal stated that the Polish Ministry of Health is seeking . . . "assistance in tailoring a robust and rigorous methodology for determining which services in Poland should be contained in the basic benefits package . . ." Working Paper 6, as indicated by its title, is meant to address this challenge.

The paper has five sections. The first and second sections outline the "robust and rigorous methodology" mentioned in the previous paragraph, in terms of an ideal process of defining a health benefits package and a system for determining coverage of health interventions. The third sections uses this methodology, or a modified version of it, to illustrate how services may be determined to be part of the benefits package, depending especially on existing evidence of efficacy/effectiveness (health benefits). This work has been carried out by the Polish Association of Quality Promotion in Krakow. The final two sections discuss implications for the future.

Contents

1 An Ideal Process of Defining Health Benefits Coverage — 7
1.1 Proposed Standard for Services to be Included in Basic Benefits Package — 8

2 A System for Determining Coverage of Health Interventions — 11
2.1 An Overview of the Coverage Process — 11
2.2 Detailed Presentation of the Proposed Coverage System — 12
2.2.1 Identification of the Technologies — 12
2.2.2 Setting Priorities Between Assessments — 18
2.2.3 Compilation of Existing Information — 19
2.2.4 Synthesis of the Literature — 24
2.2.5 The Coverage Process — 38
2.2.6 A Simplified Process of Assessment — 40
2.2.7 The Coverage Body — 42

3 Assessment Conducted by the Polish Association of Quality Promotion Concerning the Basic Benefits Package — 45
3.1 Introduction — 45
3.2 Assumptions — 45
3.3 The Structure Used in the Work — 46
3.4 Process — 50
3.5 Human Resources, Timetable and Final Outcome — 52

4 Basic Benefits Package - Implications for the Future — 53

5 Discussion — 57

6 References — 59

Appendices
A Websites Related to Health Technology Assessment

1 An Ideal Process of Defining Health Benefits Coverage

Working Paper 1 introduced the concepts of "basic benefits package" and health insurance "coverage". That material will not be repeated here.

There are several dimensions to the problem of defining the benefits package. One is that resources are limited, so services must also be limited. Another dimension is that many services provided in health care have not been shown to be of benefit, and quite a number have been shown to be harmful and probably lacking in benefit altogether. Because of the limitations that must be made, a method for determining what services will be provided and what services will not be provided is needed. As Eddy (1996) has stated,

> *"Everyone can agree that costs must be controlled but quality must not suffer. The pathway to this goal is to encourage services that have high value while discouraging the use of interventions that provide little or no value."*

Determinations of "high value" and "little or no value" can be made through the process of health technology assessment (HTA), also discussed in Working Paper 1, and developed at length in Working Papers 2-4.

A major international trend, as described in Working Paper 1, is to base coverage decisions on HTA. Eddy (1996) summarises the lessons of this experience as follows:

> *"First, it is appropriate for plans to define criteria that limit the services they will provide or pay for. That is, there is no intrinsic obligation for plans to pay for anything that anyone might want.*
>
> *Second, if plans are to accomplish these limitations, the language they use must be as precise as possible – not just legally precise, but comprehensible to members. This has several implications. An immediate one is that the unadorned words "necessary," "appropriate," and "investigational" will not work. They have been tried and they have failed. New benefit language will have to lay out specific criteria, and probably should illustrate the criteria with specific examples. Another implication is that the language should be kept as simple as possible, searching for clear lines and using common words to the greatest extent possible. Still another implication is that when there are no clear lines or when common words can be misinterpreted, specific inclusions or exclusions should be used liberally to address any services that might be ambiguous or controversial.*
>
> *The third lesson is that because the natural incentive of patients is to expand rather than contract the interpretation of the language, and because the courts will tend to interpret any ambiguity or misunderstandings in favor of the patient, if the criteria are to achieve the objective of unambiguous communication between plan, physician, and member, they should be defined narrowly. A plan can easily*

> *expand a criterion, either formally by adding specific inclusions, or informally by simply paying for services that fall outside the criterion. However, plans cannot move in the other direction; they cannot expect to win if they withhold coverage for services that fall within a liberal or even ambiguous interpretation of the criteria."*

Eddy goes on to propose coverage language based on a workshop involving insurers in the United States. The proposed coverage criteria are as follows:

1. Health plans are required to cover health interventions within the specified benefit categories if they meet the following criteria:

1.1. The intervention is used for a medical condition.

1.2. There is sufficient evidence to draw conclusions about the intervention's effects on health outcomes.

1.3. The evidence demonstrates that the intervention can be expected to produce its intended effects on health outcomes.

1.4. The intervention's expected beneficial effects on health outcomes outweigh its expected harmful effects.

1.5. The intervention is the most cost-effective method available to address the medical condition.

Section 2 is made up of definitions that are similar to those given in this report so they will not be quoted here.

Section 3 states,

"Nothing in this language prohibits health plans, at their discretion, from covering health interventions that do not meet these criteria." As will be discussed later, coverage decisions are not made solely on the basis of scientific information concerning cost-effectiveness. Health plans must not tie their hands to this sole criterion, leaving out political, cultural, geographic, and other factors. **The important point is that the procedure be transparent and understandable.**

1.1 Proposed Standard for Services to be Included in Basic Benefits Package

The central issue concerning definition of the benefit package in relation to health technology assessment, then, *is whether the intervention can be shown to have a positive effect on health outcome or not.* If not, it generally is not covered. That is the position taken in this project.

The proposed language, as noted by Eddy (1996), avoids terms such as "medically necessary", "medically appropriate", "investigational" and "safe". These terms are very difficult to define precisely, and different people have different definitions of them. If coverage criteria are to be transparent and acceptable to policy-makers, clinicians, and the general public (including patients) it is very important that the terms used can be defined rather precisely. The criteria also avoid terms such as "standard and accepted",

"widely used", and "community standard". This is because wide acceptance and use is not an accurate indication of effectiveness. *Effectiveness can only be determined definitively by scientific evidence derived from well-designed studies.*

Another point is that, as pointed out in Working Paper 1, there are many policy areas that have a relation to coverage, including regulation, mandated payment, and indications for use. It is important not to assume that coverage can solve all problems. It is an important step in itself, but it must be supported by other developments. Licensing, as laid out in Working Paper 5, is one of those. General developments in HTA and evidence-based medicine are another.

Finally, the proposed criteria mention cost-effectiveness. Cost-effectiveness must play a part in the health care system because of the limited resources for health care. The criteria concerning health benefit offer the possibility of removing interventions that are harmful or of no known benefit. However, they do not help with the problem of an intervention that is of great benefit under certain conditions of use and of little benefit under other conditions of use. The problem is that cost-effectiveness methods have not reached the stage of validity and reliability that they can be a central part of decision-making as evidence of efficacy can be. Furthermore, the available literature on cost-effectiveness is lacking, and what there is not of very high quality. Overall, the methods and results of cost-effectiveness studies have not been generally accepted. Furthermore, Poland does not have the expertise in health economics to allow in-depth consideration of cost-effectiveness. Nonetheless, costs and cost-effectiveness will inevitably play a greater and greater role in Poland in the future, so the coverage procedure must be designed to keep this fact in mind.

The final point to emphasise by way of background is that developing a procedure for improving coverage decisions does not necessarily have any relation to the method of payment of clinicians or hospitals. It is sometimes assumed that definition of the benefit package is linked to fee-for-service payment. In other words, that defining the benefit package has the purpose of controlling decisions in a fee-for-service system. Such an assumption is false. While some countries do in fact use this process as a control mechanism on physicians who are paid by fee-for service, this is not a central issue in defining the benefits package. All health care systems must define a benefit package, either implicitly or explicitly.

The primary purpose of setting standards for the benefit package is to reassure all actors in the field that decisions are fair and are oriented to the main goal, to improve health within the available resources for health care.

2 A System for Determining Coverage of Health Interventions

2.1 An Overview of the Coverage Process

The EUR-ASSESS project (Cranovsky et al, 1997) described an idealised model for coverage policy as follows:

> "*1. Identification* of the technology in question. The identification might be done by either party. The Coverage Body must determine if it is a potential priority for an assessment. If so, it would ask the HTA Body for an assessment.
>
> "*2. Literature review* to determine the availability of information. If sufficient information for an HTA is available, the HTA Body would propose an assessment to the Coverage Body, including a time frame for the assessment. If not, the HTA Body would propose supporting a prospective, well-designed study, such as a randomised clinical trial, and perhaps a simultaneous cost-effectiveness analysis to develop new information. If this proposal is accepted, a delay of several years until the final decision can be expected. Naturally, the Coverage Body may reject the option for an assessment and make the decision on other grounds. It can also, depending on the characteristics of a technology and of the clinical condition to which HTA will be applied, propose temporary conditions with coverage (as in Switzerland and the Netherlands).
>
> "*3. Synthesis of available information*, including that on efficacy, safety, efficiency, and social and ethical aspects, leads to an HTA. Extensive expert input is sought in this synthesis process. The synthesis leads to judgements, conclusions, and (perhaps) recommendations from the HTA Body or program to the Coverage Body. While Coverage Bodies obviously need clear indications concerning the value of a technology, whether recommendations are appropriate depends on the specific context. Some Coverage Bodies may prefer not to have recommendations, since they have the effect of bringing pressure for a specific action. Other Coverage Bodies may ask for recommendations to consider.
>
> "*4. A coverage proposal* is developed by the Coverage Body. After expert review, including review by the HTA Body, the coverage decision is made, published, and promulgated."

The purpose of this report is to develop the system described here in sufficient detail so that Polish authorities may develop and use the system.

This brief description obviously raises many questions.

1 What is the HTA Body?

2 What is the Coverage Body?

3 What is expert review?

In developing this paper, TNO has attempted to answer all such questions.

2.2 Detailed Presentation of the Proposed Coverage System

Figure 1 shows the proposed coverage system and its functions in the form of a flow chart. This section will explain this flow chart in some detail.

2.2.1 Identification of the Technologies

The technology to be assessed for coverage purposes must obviously be relevant for the Polish health services. In other words, the technology considered must be of concern or potential concern as a coverage issue.

Health technology may be identified by a number of means. The most usual, and perhaps most important, is that a coverage decision must be made. Generally, this means that a hospital or physician (or group of physicians) have sought payment for a certain defined service that is not presently covered. When an active coverage policy is in place, and payment will not be made without an explicit coverage decision, many issues will be identified by this obvious mechanism.

Another possibility is that the technology is of policy concern. A common story is that industry approaches a prominent clinician with a new technology relevant to his/her specialty. The clinician decides to seek coverage for a service with the new technology, since coverage may not be made available if a new machine, for example, is purchased without policy approval. The clinician in this case has several options; these options are not mutually exclusive; in some cases, all are used.

One option is to approach the policy maker directly with a plea for approval of coverage. If the policy maker has a relevant health problem and the clinician happens to be involved in the care of the policy maker, the decision may be quickly made. However, the clinicians have other possibilities. He/she can involve the media in publicising the new technology, stating, for example, that the lack of coverage means patients will miss a chance to cure in the case of a certain condition. He/she may also involve groups of patients with a certain condition in such an effort. In some cases in Western Europe, it is known that clinicians, industry, and patient groups work toward this end, with funding for the clinician and patient groups from industry. The goal is to provoke demand for the new service that the policy maker will respond to. Of course, this situation exists with or without a coverage process. In fact, a rigorous coverage process is probably the best defence against such attempts to provoke demand.

Alternatively, the policy maker may learn about a new technology that is likely to enter the market. Such knowledge could come from visits to other countries, from visitors from other countries, or from media reports. Policy makers tend to be most interested in very expensive new machine-based technologies. These technologies may not be the

most important to the health system, but the coverage process must obviously be responsive to the policy maker such as the Minister of Health.

The fact that coverage is to a fairly large extent a political process (as will be described later in this paper) means that it is beneficial if the HTA Body attempts to anticipate technologies that will be of concern. The early experience of HTA showed the importance of anticipation, for example, in the case of the computed tomography (CT) scanner. In the United States, for example, the CT scanner was not visible to national policy makers until 1975-1976, while the first CT scanner was purchased and installed in the United States in 1973 (OTA, 1978). By the time policies tried to deal with the CT scanner, it had diffused widely in the United States. Such experiences showed the importance of identifying technologies so that assessment could start early in diffusion.

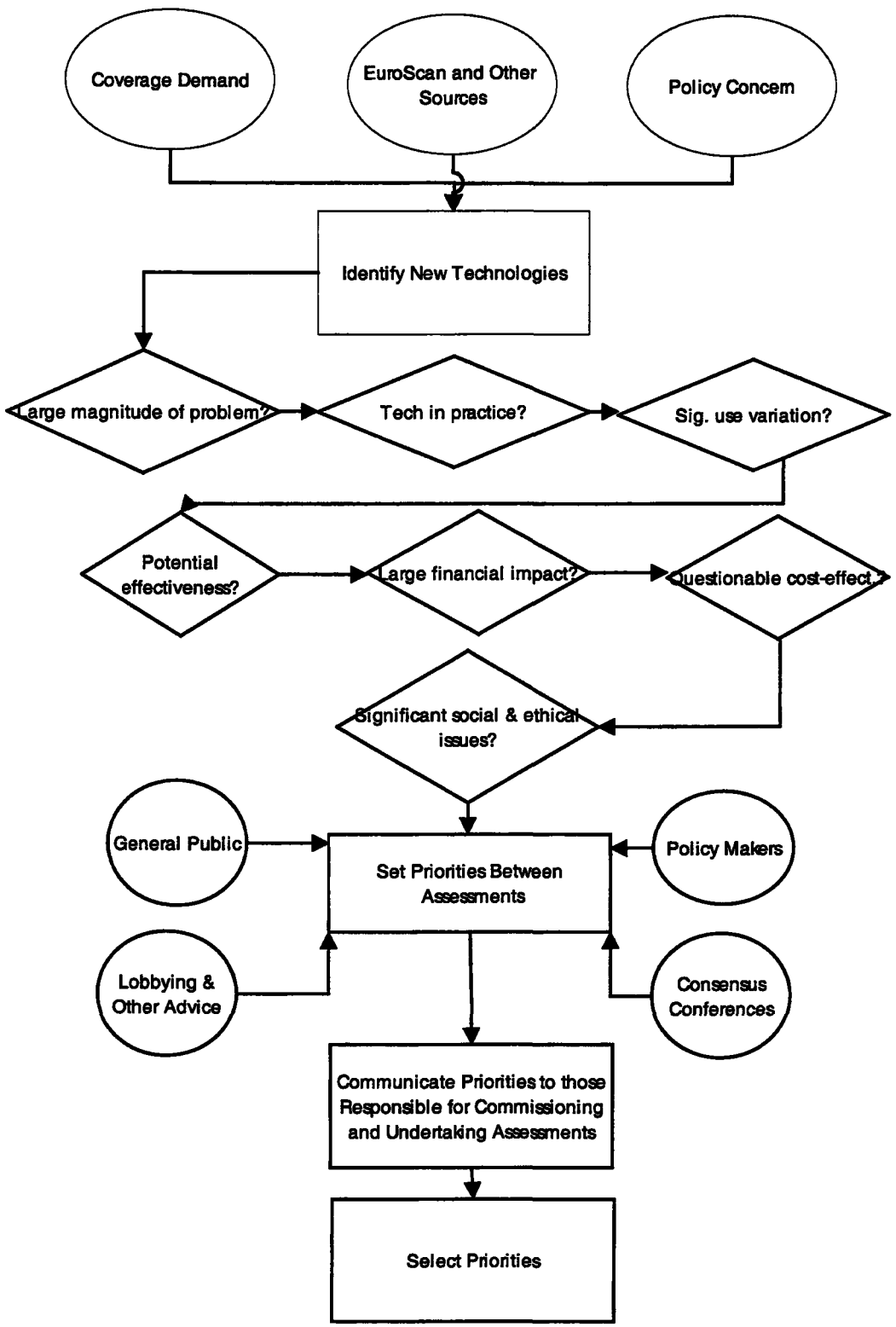
Because of this felt need, those involved in "horizon scanning" or "early warning" for new health technologies began to discuss collaboration in the mid-1990s. In 1997, The European Information Network on New and Changing Health Technologies (EuroScan) was established. EuroScan is a collaborative network of health technology assessment agencies carrying out early warning activities. The members of EuroScan aim to share and evaluate key information on selected emerging health technologies or new applications of existing ones in order to address their effects and the anticipated short and long term consequences of their use for health care and society. EuroScan also aims to support national agencies and HTA organisations in developing and running systems to provide information to health planners and policy makers on important new and changing health technologies.

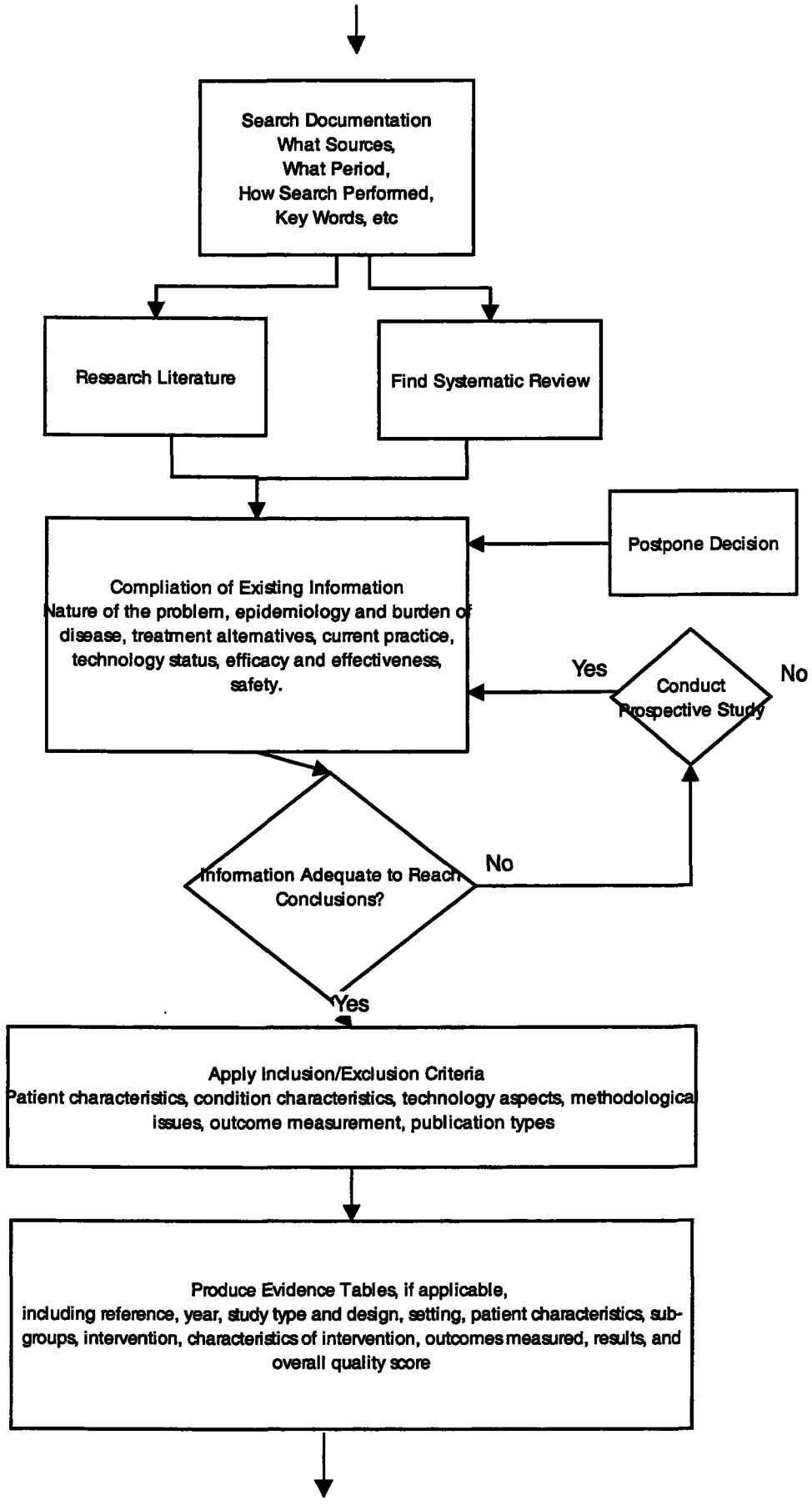
The EuroScan collaboration intends to consolidate the use of the early warning database within the work of individual agencies, promote the value of the collaboration to others with an interest in the work, and develop a research programme to evaluate activities and gain a greater understanding of the determinants of diffusion and impact of innovation in health care. Scenarios for future collaboration and information sharing have been explored by EuroScan members and range from maintenance of the 'status quo' to the institution of a formal centralised permanent international system with standardised outputs. Mechanisms for information sharing both within and external to the collaboration have also been considered.

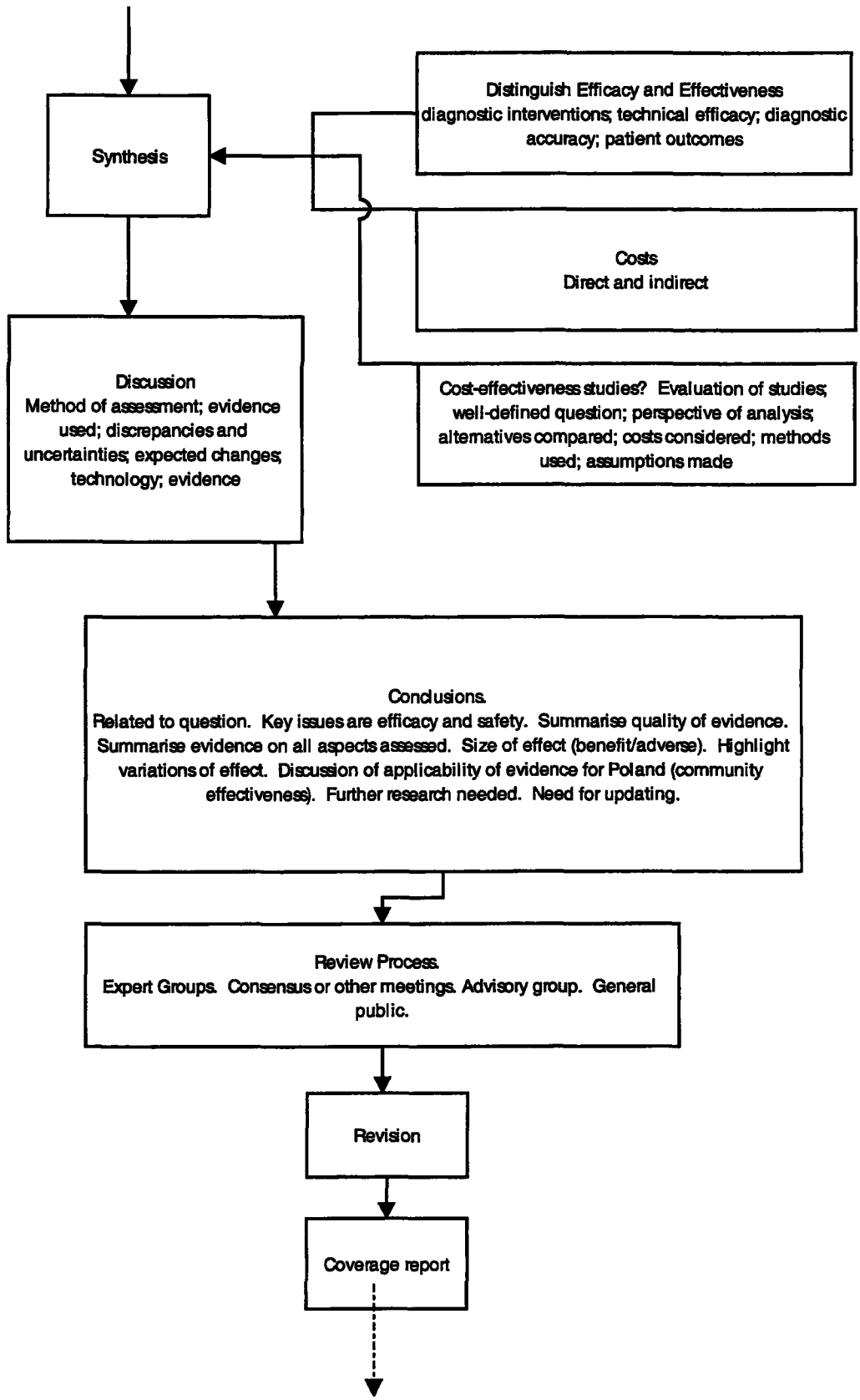
Therefore, increasingly early information on new technologies is available through EuroScan. Such information should be part of the coverage process, so that new technologies can be anticipated before they become an issue. Without such anticipation, pressures for coverage may make it impossible to carry out a thorough assessment before the coverage decision is made.

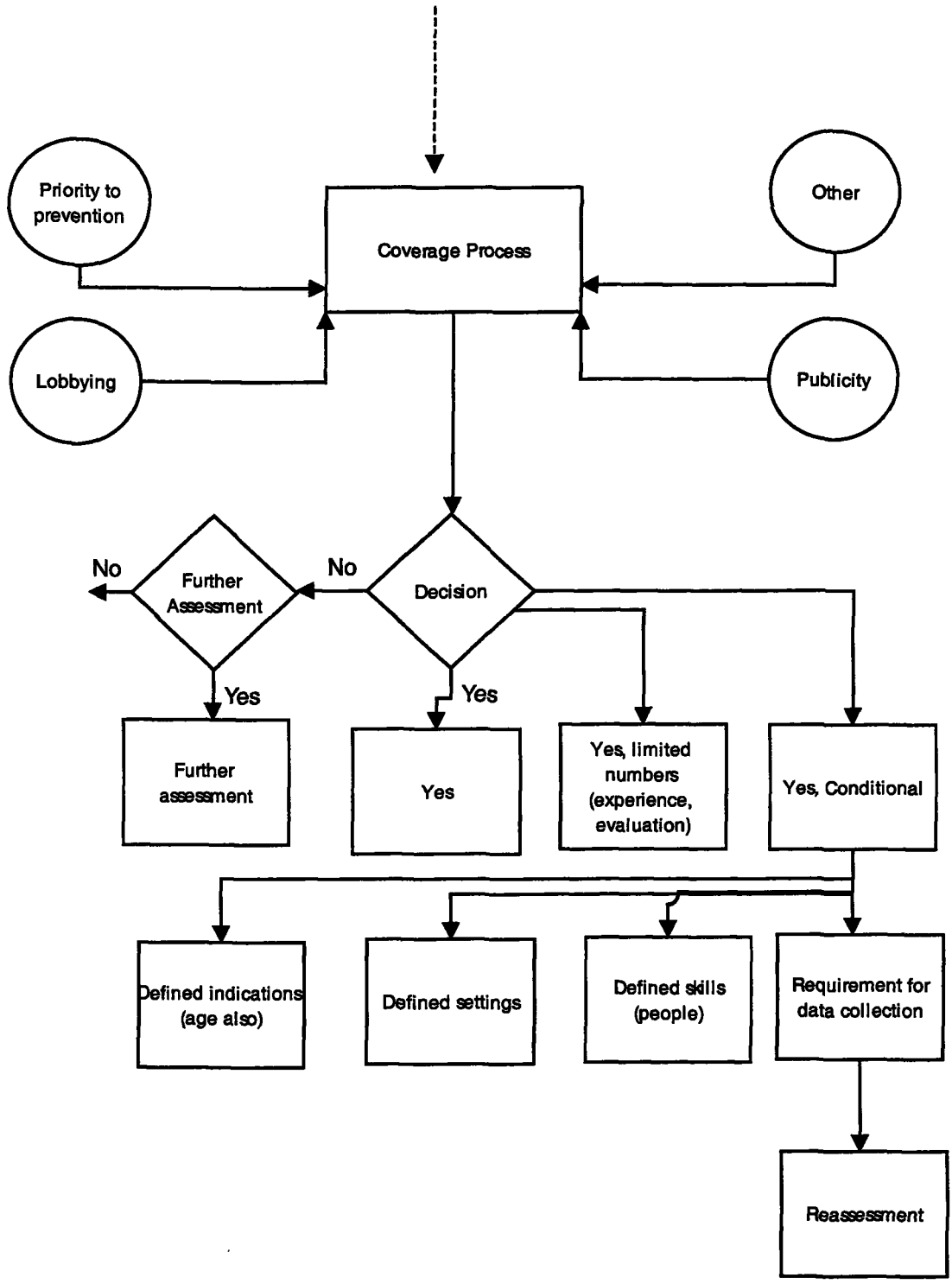
Finally, coverage is not necessarily meant to deal only with new technologies. Many older technologies have not been well-assessed, or have been found to have questionable benefits but are still paid for. An assessment can also deal with such cases for coverage decisions. In this case, the result may be a **negative list**, that is, technologies that will no longer be paid for. The identification of such technologies can be done using input from clinicians, patients or policy makers. However, here the process must be more pro-active, since the technologies do not come automatically for a coverage decision.

Figure 1: Process of Coverage Determination/ New Technologies




Search Documentation
What Sources,
What Period,
How Search Performed,
Key Words, etc
Research Literature
Find Systematic Review
Postpone Decision
Compliation of Existing Information
Nature of the problem, epidemiology and burden of disease, treatment alternatives, current practice, technology status, efficacy and effectiveness, safety.
Yes
Conduct Prospective Study
No
Information Adequate to Reach Conclusions?
No
Yes
Apply Inclusion/Exclusion Criteria
Patient characteristics, condition characteristics, technology aspects, methodological issues, outcome measurement, publication types
Produce Evidence Tables, if applicable,
including reference, year, study type and design, setting, patient characteristics, sub-groups, intervention, characteristics of intervention, outcomes measured, results, and overall quality score


Synthesis
Distinguish Efficacy and Effectiveness
diagnostic interventions; technical efficacy; diagnostic accuracy; patient outcomes
Costs
Direct and indirect
Discussion
Method of assessment; evidence used; discrepancies and uncertainties; expected changes; technology; evidence
Cost-effectiveness studies? Evaluation of studies; well-defined question; perspective of analysis; alternatives compared; costs considered; methods used; assumptions made
Conclusions
Related to question. Key issues are efficacy and safety. Summarise quality of evidence. Summarise evidence on all aspects assessed. Size of effect (benefit/adverse). Highlight variations of effect. Discussion of applicability of evidence for Poland (community effectiveness). Further research needed. Need for updating.
Review Process
Expert Groups. Consensus or other meetings. Advisory group. General public.
Revision
Coverage report


Priority to prevention
Coverage Process
Other
Lobbying
Publicity
No
Further Assessment
No
Decision
Yes
Yes
Further assessment
Yes
Yes, limited numbers (experience, evaluation)
Yes, Conditional
Defined indications (age also)
Defined settings
Defined skills (people)
Requirement for data collection
Reassessment

The Netherlands has pioneered the development of a system for assessing older technologies with the goal of removing coverage for technologies that are not of benefit. About 1995, the Dutch Sickness Funds Council asked many experts, especially physicians advising Sickness Funds on insurance issues, to propose candidates for assessment among "established" technologies. A long list of candidates resulted from this inquiry. A modified Delphi process was used to set priorities among these technologies, resulting in 126 technologies for initial evaluation. Because many of these technologies had never been carefully assessed, the process of evaluating these technologies is still ongoing in the Netherlands. In each case, the intention is to review and synthesise the literature and to remove technologies that do not appear to be of benefit. In cases where the literature is not sufficient for making such a determination, prospective research is being supported. This is obviously a very long process, but ultimately the hope is to evaluate all older technology that has not been assessed.

2.2.2 *Setting Priorities Between Assessments*

It may not be possible to carry out all assessments that would be desirable. The work of defining coverage is a large one, while assessment resources are too small for the entire task in every country. While international collaboration can assist in dealing with this problem, it may still be necessary to define priorities. In this case, an explicit and transparent priority-setting system is desirable.

In the EUR-ASSESS project (1994-1997) the subject of priority setting for HTA was studied. In the report on priority setting an analysis of the process of setting priorities is presented. The report offers some practical guidance to help those wishing to set priorities for HTA to develop a system suited to their particular needs and circumstances (Henshall et al, 1997). The subject of priority setting for HTA is further elaborated in Working Group 2 "Developing systems for the routine exchange of information between programs" of the ECHTA project, which was completed for presentation to the European Commission on 15 October 2001. That report will also be available within a few months.

Priorities are set using rather obvious criteria, such as potential benefits of the technology (which includes the number of people who may benefit) and potential costs of the technology. It would seem desirable to include social criteria as well as such more quantitative criteria such as effectiveness and financial costs. However, it is uncommon for social issues to be explicitly considered in priority setting.

The main point about setting priorities is that the main question concerns the benefits that will result from an **assessment**, not necessarily from the technology in question. In other words, if the decision has already in effect been made, for example, there may be little reason to carry out an assessment even though assessing the technology would seem to be important. In the case of coverage, this may not be much of a factor, since presumably assessments should always have some effect on decision-making.

After priorities for assessment are selected, the assessment process can begin. It should be noted that the body identifying technologies and setting priorities for assessment might not necessarily be the same as that carrying out the assessment. The system for examining coverage issues and actually making coverage decisions could be in several different organisations. For example, priorities might be defined by a special part of the Ministry of Health, which might have better access to relevant information than an HTA Body. However, HTA should be carried out in an HTA program that is relatively

removed from political factors. Otherwise, HTA will be seen as political, which will make the results less acceptable to different parties. This point will be further emphasised later concerning the separation of production of the coverage report from the actual coverage decision.

2.2.3 *Compilation of Existing Information*

The policy question is now clear. The Coverage Body or policy maker must make a coverage decision in a specific case. It is the task of the HTA Body to produce information that will be useful for this decision.

The HTA Body will begin with the collection of background information. In an ordinary HTA, it is desirable to answer a series of questions before beginning:

1. Who initiated the report?
2. Why is an assessment needed right now?
3. What type of decision is the HTA intended to influence or support?
4. What is the primary audience for the report?

In the case of a coverage decision, however, the answers to these questions should be clear. The process has generally been initiated because of a coverage concern and the assessment it needed for that purpose. The primary audience is the Coverage Body.

Other background information needed includes the following:

1. Nature of the health problem
2. Epidemiology and burden of disease
3. Treatment (or diagnostic) alternatives
4. Current practice
5. Technology status
6. Efficacy/effectiveness
7. Safety

Sources for such information include:

1. Research literature
2. Routinely collected data (health status, costs, utilisation)
3. Available clinical guidelines
4. Special sources (disease registers, organisations of affected people, experts, manufacturers

5. Other HTA reports

Box 1 gives questions to be addressed in seeking information on the health condition and the target group.

Box 1. Questions to be addressed as background information on condition and target group (Adapted from DES; Burls et al. 2000)

Questions	Example
• **Condition(s)**	• Health problem • Disease
What are the mechanisms of disease?	• Causes • Pathology
What is the course and prognosis of the condition?	• Clinical presentation • Stages • Time course
What are the consequences? (Outcomes)	• Physical disabling • Psychological consequences • Death
Treatment alternatives and current practice	• Drugs • Surgical • Current service provision
• **Target group(s) (epidemiology, burden of disease)**	• Patients • Healthy subjects (for prevention)
• How many people are affected?	• Incidence • Prevalence
• Who is affected?	• Age • Gender • Social factors • Risk factors

Source: Busse et al, 2001.

It is also desirable to give background information on the technology. The technology could be a drug, a device, a community-intervention, a medical aid, a medical or surgical procedure, an organisational process, or possibly a combination of some of these. Questions to be addressed are presented in Box 2.

While this information is important in assessing the technology, the next step, seeking information on **efficacy and safety**, is the core of the process. Existing information may also give data or insights on psychological, social, ethical, organisational, professional, and economic aspects of the technology. Box 3 gives examples of outcomes for these different areas. Box 4 outlines the general steps for addressing each aspect of the assessment. The first task is to find out the extent to which important questions have been addressed by others. If syntheses have not been carried out by others, a thorough examination of existing literature will be necessary.

For different aspects of the assessment, different sources of information may be useful. Probably the most useful first step is to determine if the questions – or some aspects of the questions – have been examined by others, especially HTA agencies. Annex 1 gives some handy sources of information available on the Internet. A survey as part of the ECHTA project found that those working in HTA found the HTA database the most useful starting point in an assessment. For the purposes of coverage, an HTA report done in another country may furnish sufficient information. However, the report must still be examined for quality, relevance, timeliness, etc. It cannot be assumed that any

source automatically gives information that is reliable and valid. Quality criteria will be further discussed below.

Box 2. Questions to be addressed as background information on the technology	
Question	Aspects / examples
• How does it work? What kind of intervention is it?	• If a device, explain technical characteristics, functioning • If a community/system related intervention, explain its crucial features
• What are the requirements for its use?	• Setting for use/implementation • Special measures needed for use/implementation • Qualification required • Maintenance
• What is the status of the technology?	• Diffusion/distribution • Patterns of use • Current indications for use • Current utilisation • Costs • Regulatory status • Manufacturers and market shares

Source: Busse et al, 2001.

If primary research information is necessary, a standard method of searching should be used, such as the one developed by the Cochrane Collaboration (see Annex 1 for the address of the Cochrane website). A systematic approach can be used for all aspects of the assessment, and not only efficacy. The key point is that the appropriate literature be searched. The Medline system of the National Library of Medicine is probably most appropriate for information on efficacy. Other databases will furnish data on other aspects, such as cost-effectiveness or ethics. Documentation of the information sources is very important for transparency of the process. Both literature used and literature not used should be included in the documentation (Box 5).

Inclusion/exclusion criteria need to be defined before the search is undertaken to assure transparency (Box 6). These criteria need to be defined in advance. This will avoid accusations that those searching the literature have excluded some literature in order not to report certain findings.

Box 3. Examples of outcomes for different aspects of HTA	
Aspect of assessment	Outcomes
Safety	• Mortality directly related to the use of technology • Morbidity/disability directly related to the use of technology
Efficacy/Effectiveness	• Change in overall/ condition-specific mortality • Change in morbidity/ disability/ disease-free interval • Change in quality of life • Change in quality-/disability-adjusted life years (QALYs/DALYs)
Psychological/ Social/ Ethical	• Compliance • Acceptance • Satisfaction • Demand • Preferences • Information/patient advice requirements
Organisational/ Professional	• Utilisation of service • Change in the treatment location • Change in length of hospital stay • Change in required personnel, material inputs (e.g. hospital beds) and organisational structure • Training requirements
Economic	• Costs and changes in cost compared to current practice (if applicable) • Cost-effectiveness, cost-utility, cost-benefit

Source: Busse et al, 2001

Box 4. General methodological steps for addressing each aspect of assessment
• Searching for sources of information
• Selecting and evaluating information (application of inclusion and exclusion criteria)/ appraising the evidence
• Synthesising the obtained data

Source: Busse et al, 2001.

Box 5. Documentation of the sources (DIHTA; Kristensen et al. 2001)
• Which sources have been consulted?
• Which period did the performed search cover?
• How was the search performed? (Strategies, key words, search criteria)
• When was the search conducted?

Source: Busse et al, 2001

If primary research information is necessary, a standard method of searching should be used, such as the one developed by the Cochrane Collaboration (see Appendix A for the address of the Cochrane website). A systematic approach can be used for all aspects of the assessment, and not only efficacy. The key point is that the appropriate literature be searched. The Medline system of the National Library of Medicine is probably most appropriate for information on efficacy. Other databases will furnish data on other aspects, such as cost-effectiveness or ethics. Documentation of the information sources

is very important for transparency of the process. Both literature used and literature not used should be included in the documentation (Box 5).

Inclusion/exclusion criteria need to be defined before the search is undertaken to assure transparency (Box 6). These criteria need to be defined in advance to avoid the charge that those searching have excluded some literature because if might have findings not deemed desirable by those carrying out the assessment.

Box 6. Issues addressed in inclusion and exclusion criteria

- Patient characteristics (e.g. age, gender)
- Condition characteristics (e.g. stage of disease)
- Technology aspects
- Methodological issues (e.g. number of patients, length of follow-up, study design)
- Outcomes measured
- Publication type

Source: Busse et al, 2001.

An especially important part of the inclusion/exclusion criteria is the study design. Hierarchies of study design have been developed, to be discussed further later. It is usually considered that randomised trials (RCTs) or meta-analysis from RCTs are the highest level of evidence. If such studies are available, other studies designs may be excluded from the analysis. However, for some aspects of the technology, such as social or ethical considerations, such a hierarchical system is not relevant. Furthermore, one should not necessarily conclude that any RCT presents valid information. The RCTs identified must themselves be assessed for quality. A number of checklists are available for this purpose.

In defining the search strategy, the searcher must decide the period of the search, strategies of the search, use of key words, search criteria, and how the search is actually to be conducted. Searching the literature using electronic databases has become a highly specialised procedure. Ideally, those carrying out searches should have access to a trained documentalist or librarian.

When the existing information is inadequate to reach a conclusion. While existing information can often answer the main questions of the assessment, particularly efficacy, it is often so that the information is inadequate. In that case, there are only two feasible strategies, both involving delaying the assessment.

One is to organise or fund one or more prospective studies to answer the questions of the assessment. Alternatively, one can recommend that other bodies fund such research. This is frequently done in the Netherlands and the United Kingdom, which have funds for this purpose. The prospective research is not carried out by the HTA Body, but generally by academic groups that have the clinical, epidemiological, and possibly the economic expertise to organise the needed studies. In some cases, prospective analyses of other issues, such as ethics, may be funded for such a purpose.

The other possibility is to postpone the assessment, hoping that more information will become available through research or assessment in (for example) other countries.

A problem that may arise is that pressure for a decision may be very great. In such a case the policy-maker or the Coverage Body may not be able to postpone its own decision. However, it is of the utmost importance that the HTA Body not present a report based on poor evidence that recommends coverage of the technology. If the Coverage Body wishes an assessment of all the medical literature, that can be done, but it should make the poor quality of the evidence clear. Likewise, the Coverage Body may wish to learn what a consensus of the medical profession or of a group of specialists thinks about the issue. This may also be a legitimate part of an HTA process, but it still needs to be clear that the evidence is poor.

In a later section, the presentation of recommendations based on different levels of evidence is presented. This is one approach to this dilemma.

In the Polish context, it is probably unlikely that much prospective research can be commissioned or carried out for some years. Therefore, postponing the assessment or basing it on sub-optimal evidence may be the only alternatives in many cases.

The evidence is adequate to reach conclusions. As described above, the literature has been searched to determine if it is adequate using inclusion/exclusion criteria. As a first step to synthesis of this literature, it may be desirable to develop evidence tables as a further indication that the process is rigorous and transparent. The Figure shows the types of information that should be included in the evidence tables. Evidence tables may not be necessary if there have been few RCTs, which is usually the case in non-drug areas of assessment. With more studies, evidence tables are an aid to the assessor, but they should also be included in the report.

2.2.4 *Synthesis of the Literature*

The first step is a qualitative synthesis of the literature. As already noted, evidence tables may be very useful at this stage of the assessment. All relevant information should now be pulled together. After this qualitative synthesis, a decision can be made if a quantitative synthesis, usually referred to an "meta-analysis", should be carried out. A meta-analysis as a mathematical calculation requires considerable expertise, so it will not be discussed here. However, good sources discussing meta-analysis are available. One example is Petitti (1994).

There are a number of specific methodological issues that almost always arise in carrying out the literature synthesis. Assessing **safety** implies a wide scope of searching for possible harm caused by the technology. The sources of information include the published literature and routinely collected data, such as that from regulatory authorities. A severe problem in assessing safety is that RCTs are not generally an adequate source of information because adverse effects often occur in too small rates to be see in an RCT, and also often occur some time later, while RCTs are time-limited, often to three years. Furthermore, reporting of safety in RCTs is sometimes not done adequately. Therefore, other types of studies, such as observational studies, may be more useful in assessing safety. The different sources of information on safety need to be documented, preferably in a table in the report. Important aspects of safety should be discussed, such as whether the problem is with a device or with an operator of the device, and what the time frame for clinical harms may be. It is often difficult to determine rates of different harmful effects, so a list of effects may be the only solution.

Efficacy and **effectiveness** of a health technology both refer to the health benefit to be obtained from a health technology. **Efficacy** refers to use of the technology under ideal circumstances, such as conditions of a well-controlled study in an excellent medical centre. **Effectiveness** is the extent to which the technology works in ordinary day-to-day practice. Effectiveness is probably always less than efficacy, that is, less than expected. The accepted method for assessing efficacy is through RCTs or syntheses of RCTs. However, RCTs are often not feasible for different reasons, requiring other study designs. Furthermore, RCTs recruit people for specific reasons while excluding others. For example, RCTs involving children are relatively rare because children cannot give informed consent. RCTs involving women are less common than those involving men for a number of reasons, including the possibility that the woman may be pregnant. Thus, the effects of the technology from an RCT may not be fully applicable to the "real world".

If data from RCTs are not available, there are often data from other types of studies. Controlled clinical trials are often done with non-randomised designs. While the control group in such a case is not strictly comparable to the study group, such studies may furnish valuable information. RCT data are better, but presumably any evidence has some usefulness. Other study designs, such as case control studies, may also give useful information. Given the pressures on the assessment process inevitable in an assessment done for coverage purposes, a good rule of thumb might be to first seek RCTs (or even better, systematic reviews of RCT's). Failing RCTs, a search for other types of evidence might be made. A significant problem here, however, is that there is no consensus on how evidence from other studies can be synthesised, or how evidence from such studies can be synthesises with evidence from RCTs.

Systematic reviews or syntheses are now often available and are a great help in assessing a technology. As already mentioned, there are agreed-upon principles for carrying out a systematic review. However, the quality of systematic reviews must be considered. Box 7 gives some key issues in assessing systematic reviews. It is obvious that these criteria are similar to those that should be used in carrying out a synthesis or assessment.

Box 7. Key issues in assessing systematic review articles

- What are the review questions? Are they relevant for the current research questions?
- Which sources were searched? How were they searched?
- Are selection criteria explicit and appropriate?
- What criteria were used to assess study quality?
- How were the data extracted?
- How were the data synthesised?
- Are the results of the review transferable to my context?
- Should the review be updated?

Source: Busse et al, 2001

In contrast to reviewing the evidence on efficacy from RCTs, there is little consensus in how to measure effectiveness, or what is often called "community effectiveness." A number of factors go into determining effectiveness, including efficacy, diagnostic accurate, professional compliance with good practice (including training), patient

compliance with the professional prescription or advice, and coverage of the technology. Effectiveness is thus a combination of efficacy with system-, provider- and patient-related variables. Since many of these variables are discussed as part of the assessment, and because there is no consensus on how to measure effectiveness, these variables need special attention in the appropriate section of the report.

There is a strong consensus that efficacy should be measured using **health outcome** information, such as mortality or morbidity. Use of other outcome measures, often called **surrogate outcomes**, is a common practice in the literature. Examples of surrogate outcomes include biochemical and physiological outcomes. For the analyst, these outcomes must be approached with scepticism. The main rule is that surrogate outcomes cannot be considered as strong evidence unless they have been empirically connected to health outcomes in high quality studies.

Assessing **diagnostic technology** raises particular problems (see Box 8). Historically, evaluation of diagnostic technology has involved "technical efficacy", that is, technical performance of a test. A good example is the quality of the image with an x-ray or other diagnostic imaging device. However, the rule for efficacy here should be that the diagnostic test is related to health outcomes. More and more often, studies of diagnostic technologies involve patient health outcomes. Nonetheless, in many cases, the diagnostic test is today evaluated using "surrogate measures" such as sensitivity or specificity. At least, a new diagnostic test should be compared on such measures with other tests aimed at the same diagnosis. While there are proposals to improve the assessment of diagnostic tests using similar measures, in all cases the analyst should be considering if the test can lead to an improvement in patient outcome. This often involves assessment of available treatments and not only diagnosis as an outcome.

Box 8. Evaluation of efficacy and effectiveness for diagnostic interventions (Adapted from Fryback & Thornbury 1991, Flynn & Adams 1996)

"Level"	*Typical measures*
• Technical efficacy	• Physical parameters describing technical performance of the test (e.g. image quality)
• Diagnostic accuracy efficacy	• Sensitivity (% of positives among ill) • Specificity (% of negatives among healthy) • Accuracy (% of correct diagnoses) • Likelihood ratio (likelihood for a given test result in a patient with the target disorder compared to the likelihood of the same result in a patient without the target disorder; details at http://cebm.jr2.ox.ac.uk/docs/likerats.html)
• Diagnostic thinking efficacy/ effectiveness	• Post-test odds/ probability compared to pre-test odds/ probability in target population • % of cases in which test is judged "helpful" to making diagnosis
• Therapeutic effectiveness	• % of cases in which test is judged "helpful" in planning therapy • % of therapeutic procedures avoided due to test information
• Health-related effectiveness (Patient outcomes)	• Mortality/morbidity avoided with test • Changes in quality of life through use of test

Source: Busse et al, 2001.

Diagnostic tests are also often used for **screening**. Screening is the application of a test to detect a potential disease or condition in a person who has no known signs or symptoms of that disease or condition. The goal of screening is the early detection of disease or risk factors of disease so that intervention can reduce morbidity and mortality from the involved disease. In contrast to diagnosis, screening comes with a strong ethical imperative to help the health status of a patient or a population. Screening is not sought as diagnosis is, but provided by national screening programs or by clinicians in unorganised "opportunistic screening", when the patient comes for another purpose. The problem is that physicians often assume that finding a condition by screening is in itself a good outcome, while standards of public health state that screening should only be offered when the health status will be improved. This means, in particular, that there must be a proven effective intervention for the condition being sought. Efficacy of screening, then, should always be based on good studies that use health outcomes as their end measure.

Preventive interventions are aimed at avoiding the appearance of a condition in the target group. They may be implemented on an individual level, as is often the case with vaccines, or at the community level, which is more common with organised screening programs. The outcome sought here is also improved health outcome, and that should be assessed by the HTA analyst. Common problems are the long follow-up that may be necessary, the use of large observation units (communities, eg), and the difficulty in connecting outcomes to the intervention.

Psychological, ethical and social dimensions of the technology should ordinarily be assessed. The difficulty in these areas is that there is no consensus on methods to carry out the assessment. Transparency of any process used and description of results as fully as possible is then called for. Much qualitative research is in fact carried out concerning health technology, but no hierarchy has been proposed that would allow one to determine "best practice".

Organisational and professional implications may be an important part of an assessment, covering changes that may be induced by the technology and their consequences. Organisational changes that can be assessed include:

1. Utilisation of service (a new drug may lead to less surgery)
2. Changes in the treatment location (e.g. minimally invasive surgery)
3. Training/qualification requirements (the new technology may require an expert trained in its application)
4. Channels of cooperation/community (moving services out of the hospital may require new channels of community between general practitioners and specialists)
5. Job satisfaction (a new technology may require a high throughput, which may affect staff satisfaction because of limited time with patients).

Methods for collecting such information include interviews, questionnaires, and focus or consensus groups. Recommendations of manufacturers may also be relevant.

Assessment of **economic issues** implies collecting information on resource consumption from use of the technology. Further, it requires comparing costs to other outcomes, such as efficacy or effectiveness. Finally, this calculation of costs and benefits must be compared to other technologies. The purpose is to discover if the proposed investment is a wise use of resources or not.

Health economics is a large field and cannot be summarised briefly. The reports for Working Papers 2, 3 & 4 have already considered economics in some detail. Only a few aspects will be touched upon here.

Most economic analyses focus on so-called **direct costs**, that is, costs that are directly related to a health condition and the system response to that condition. Such analyses often ignore other significant costs that may not fall within the health sector, such as implications of an illness for a person's employment. The division of costs into direct and indirect costs seems without purpose. The point is that all significant costs should be taken into account. It may be that some costs are very difficult to measure. These need, in any case, to be noted.

There are a number of types of economic analysis. These are summarised in Box 9. From the standpoint of coverage decisions, cost minimisation analysis may be often the most relevant. In this type of analysis, effects are equivalent – or are assumed to be equivalent-- for two or more technologies. In this case, the less expensive technology would be favoured. Such an analysis is particularly relevant in a setting such as that of Poland, where expertise to do "higher level" economic analyses may not be available. Otherwise, some form of cost-effectiveness analysis is called for. In a coverage decision, this may be no more than an array of benefits and an array of costs.

Whatever the particular type of economic analysis, there are now commonly agreed standards for carrying out studies, as well as interpreting them. A common version is presented in Box 10. The relevance in this case is particularly in the interpretation of available studies. Unfortunately, relatively few economic studies are satisfactory on even these important parameters. An important drawback in attempting to integrate economic analysis into HTA or coverage decisions is the poor quality of most of the available literature.

Attention to all these matters leads to what has been called a "comprehensive assessment". It is important to be as comprehensive as possible, given time and resources available for the assessment. In many cases it will be necessary to focus on efficacy and safety and consider costs to some extent. In other cases, however, other dimensions may be as important as efficacy. For example, in the area of genetics, ethics plays an ever more important part, and should be considered explicitly in any assessment done in this area.

Box 9. Types of economic analysis	
Type of economic analysis	When should the specific type of analysis be chosen?
• Cost-minimisation analysis	• If the compared technologies are equally effective, then it is only necessary to collect data about costs
• Cost-effectiveness analysis	• If the effectiveness of the compared technologies are different (e.g. the difference in costs have to be weighted against the difference in effectiveness) • If activities with the same aim and measure of effectiveness are compared
• Cost-utility analysis	• If health-related quality of life is an important health outcome • If activities across specialities or departments in the health care sector have to be compared
• Cost-benefit analysis	• If non-health effects also are of importance (e.g. the treatment process itself, utility of information) • If only one technology is assessed (net-benefit) • If there is a wish that individual lives are valued in monetary units • If activities across society have to be compared

Source: Busse et al, 2001.

Whether or not the original literature is collected by the assessors, available systematic reviews will undoubtedly be used in making the coverage report. Such reviews must be examined for quality. Box 11 gives a scheme for evaluating the quality of a **systematic review**, including an HTA report published by an HTA agency or program. Another helpful source of information may be **clinical guidelines** developed in different countries. Such guidelines are also often available through internet sources (see Annex 1). Box 12 gives some indications concerning how guidelines may be evaluated for quality. The key point in both cases is whether the material is evidence-based; that is, did it result from a thorough and systematic review of available literature. This point brings out the importance of describing the method of systematic review so that the user of the report, as well as others, can determine its quality.

Box 10. Economic evaluation

- **Study frame:** clearly stated research question, identification of target population, explanation of choices and assumptions made etc.
- **Analytical technique:** choice to be explained
- **Study perspective:** societal perspective if the study does not require a narrower perspective
- **Selection of alternatives:** description and justification of choice; recommendation to use currently most effective or efficient alternative
- **Data collection:** to be described in detail; must include systematic review of literature; various types of studies and data sources are suitable
- **Costing:** all relevant direct and indirect costs should be identified, collected and reported; physical units should be reported separately from costs of resources; use of average values only if marginal data are not available
- **Outcome measurement:** primary outcome measures to be reported clearly; if values for health states are used, individual utilities should be distinct from modelling society's valuation
- **Time frame:** long enough to capture all effects; modelling can be used to estimated long-term costs and outcomes if real data are unavailable; shortening of time horizon has to be justified and possible bias estimated
- **Discounting**: necessary if costs and consequences occur at different times; use of standard rate (5%) plus national recommendation
- **Sensitivity analysis:** should be conducted to test robustness of results to a variation of assumptions, cost and outcome parameters and discounting rate
- **Equity:** values and preferences important but more valid indicators are needed

Source: Busse et al, 2001.

Box 11. Proposal for a Checklist/ Criteria for the assessment of the quality of HTA reports		
Criterion		**Questions**
A	*Basic information*	• Are the authors of the report stated? • Is/Are any possible conflict(s) of interest stated? • Is there any information about who financed the report? • Was the report externally reviewed?
B	*General methodological aspects of the assessment*	• Was there a stated HTA report protocol? Was it followed, if not why not? • Is the scope of the assessment specified? Is there an explanation given for aspects not being assessed? • Are there clear research questions posed? • Are sources of information used for each aspect stated? Is it described how was the information for the different aspects gathered? • Are selection criteria for the different kinds of information used stated? • Are validity/quality criteria for appraisal of information clearly stated for each aspect? • Were evidence tables used?
C	*Description of the context of the assessment*	• Is the reason why the HTA was conducted stated? • Is the timing of the HTA explained (e.g. inappropriate extension of indication)? • Is what decision(s) the HTA is intended to support stated? • Is there any information given of who has commissioned the HTA?
D	*Background information*	• Were conditions, target group, relevant interventions or comparisons between interventions and relevant outcomes appropriately defined?
E	*Data about the status quo of the technology*	• Are patterns of utilisation, diffusion, indications, time trends adequately described? • Is an analysis of the regulatory status of the technology provided (e.g. market admission, status in other countries)?
F	*Technical description of the technology*	• Is there any consideration of when and how technical characteristics affect the outcomes? • Description of additional influencing factors (e.g. qualification requirements of staff, quality assurance, risks)?
G	*Safety*	• Are sources of data stated? • Are selection criteria for material stated? • Is there a transparent assessment of validity/quality of data? • Are the results transparently presented?

Box 11. Proposal for a Checklist/ Criteria for the assessment of the quality of HTA reports		
Criterion		**Questions**
H	*Efficacy / effectiveness*	• Is the literature search done in a systematic way and documented accordingly (including search strategies, data sources and years)? • Are inclusion / exclusion criteria for primary studies defined? • Are included studies checked for quality and validity? • Is there a description of data extraction of included studies? • Is there a listing of excluded studies with reasons for exclusion given? • Are the results properly documented (e.g. tables, graphs, meta-analysis plots)? • Do the conclusions match the results?
I	*Psychological, social, and ethical considerations*	• Are psychological/social/ethical implications of the technology under consideration adequately discussed? • Are sources of data stated? • Are selection criteria for material stated? • Is there a transparent assessment of validity/quality of data? • Are the results transparently presented? • Are assumptions made, clearly stated?
J	*Organisational and professional implications*	• Were organisational and regulatory issues discussed (e.g. responsibility, necessary investments, financing, regulation, personnel, need, demand)? • Are the methods used for assessing these aspects stated?
K	*Economic evaluation*	• Is there a proper documentation of the methods used (see above)? • Is the perspective of the economic evaluation clarified (e.g. social insurance, societal)? • Are assumptions (e.g. for discounting rates, sensitivity analysis) justified? • Are issues of transferability (e.g. prices, cost structures, remuneration) across countries or settings adequately discussed?
L	*Discussion of generalisability / applicability of the findings*	• Are aspects of the generalisability of the results discussed (e.g. for populations not included in clinical trials or in different settings)? • Are aspects of the transferability of the results to different settings discussed (with regard to epidemiology, diffusion, structure of health care delivery, reimbursement, access)?

Source: Busse et al, 2001.

Box 12. Checklist for clinical practice guidelines

I. Are the recommendations valid?

Primary Guides:

- Were all important options and outcomes included?
- Was an explicit and sensible process used to identify, select, and combine evidence?

Secondary Guides:

- Was an explicit and sensible process used to consider the relative value of different outcomes?
- Is the guideline likely to account for important recent developments?
- ***Has the guideline been subjected to peer review and testing?***
- Were the results similar from study to study?

II. What are the recommendations?

- Are practical, clinically important, recommendations made?
- How strong are the recommendations?
- What is the impact of uncertainty associated with the evidence and values used in the guidelines?

III. Will the recommendations help in the clinical practice?

- Is the primary objective of the guideline consistent with your objectives?
- Are the recommendations applicable to your patients?

Source: Busse et al, 2001.

Discussion of methods and results. Box 13 shows the issues that should be covered in the discussion section of the report. Some of the methods used may have already been discussed, but they now need to be discussed in terms of why certain choices were made and what the implications of those choices might be.

Box 13. Discussion

- Methodology of the assessment
- Evidence used (quality, validity, generalisability)
- Assumptions made
- Discrepancies and uncertainties identified
- Expected changes (in technology, in evidence)

Source: Busse et al, 2001.

The evidence is seldom entirely satisfactory. Therefore, the assessment should discuss the shortcomings of the evidence, particularly those that may be related to the validity of the conclusions. Because of limitations of the data, assumptions may have been necessary. These should also be discussed, along with their possible effects on the conclusions. Furthermore, any uncertainties should be identified and discussed.

Finally, any anticipated changes in the technology or important aspects related to the technology should be discussed. New technology may be expected, or modifications in the old technology. Important studies may be underway that could affect the

conclusions. In addition, updating of the report will probably be necessary at some time. A rough statement of an estimated time for updating could be stated here.

Conclusions. Box 14 gives some areas that should be part of the conclusion section. The conclusion section is intended to provide answers to the policy maker. This section should summarise the most important findings of the report.

Box 14. Conclusions

- Related primarily to efficacy and safety
- Summarise quality/origin of the evidence
- Summarise evidence on all aspects assessed
- Give size of effect (benefit/adverse)
- Highlight differences among groups of patients (if found)
- Highlight variations of effect with varying characteristics of technology (if found)
- Discuss applicability of evidence for national/local context and "community effectiveness"
- Point out fields where further research is needed

Source: Busse et al, 2001.

The most important conclusion of course concerns the efficacy/effectiveness of the technology. If no efficacy has been demonstrated, the conclusion may be relatively simple. Inefficacious technology should ideally not be part of the health care system. Likewise, technology with significant safety problems and no demonstrated efficacy should not be covered. Technology with good evidence of significant efficacy is obviously of more interest than technology whose effects are less, especially when there are already alternatives available with a similar level of efficacy.

An important consideration related to efficacy is effects among different groups of patients. Technology is often tested in patients with severe disease. Once the technology becomes available, it is often used in patients with less severe disease, or even totally different conditions. The assessment should make it clear where the greatest benefit would be expected, and areas where less benefit, or no benefit, might be expected.

Since the evidence presented will come largely from non-Polish sources, the applicability of this evidence to the Polish situation needs discussion. Community effectiveness relates very much to local circumstances. Can specific situations be identified that may lead to community effectiveness that will be less than that seen in efficacy studies? Perhaps coverage should not be considered if certain conditions cannot be met.

The uncertainties and problems with the evidence lead to a statement of the need for further research. Perhaps funds are available in Poland for such research. At any rate, seeking such evidence is an important part of updating. A statement concerning a possible appropriate time for updating should be made.

Recommendations. Most likely, the policy maker or policy body would wish to receive one or more recommendations from the HTA Body. In the case of strong evidence for

efficacy and moderate costs, the recommendation is not difficult. In many other cases, however, making recommendations is not so easy.

The problems and discrepancies in available literature have led different groups to recommend that recommendations be related to the excellence of the evidence. Then, the policy maker will be able to understand better the assumptions that lie behind a particular recommendation.

Boxes 15-18 present commonly used schemes that link the quality of the evidence to recommendations. The quality of the evidence in this case, as shown in Box 15, is based on the source of the evidence, particularly in terms of the methods of studies used. In short, strong evidence comes from good randomised controlled trials (RCTs) at one end of the spectrum, while opinion of respected experts is weak evidence.

Box 15. "Traditional" Evidence-Based Medicine hierarchy of research design/quality of evidence	
I:	Evidence obtained from at least one properly randomised controlled trial.
II-1:	Evidence obtained from well-designed controlled trials without randomisation.
II-2:	Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one centre or research group.
II-3:	Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence.
III:	Opinions of respected authorities, based on clinical experience; descriptive studies and case reports; or reports of expert committees.

Source: Busse et al, 2001.

A recommendation based on strong evidence, as shown in Box 16, is then an "A" recommendation. A similar scheme, but with more details, is shown in Box 17. A scheme currently used in Poland as part of practice guidelines is set forth in Box 19.

Box 16. Grades of recommendations	
A	consistent level 1 studies
B	consistent level 2 or 3 studies ***or*** extrapolations from level 1 studies
C	Level 4 studies ***or*** extrapolations from level 2 or 3 studies
D	Level 5 evidence ***or*** troublingly inconsistent or inconclusive studies of any level

"Extrapolations" are where data is used in a situation, which has potentially clinically important differences than the original study situation.
Source: Centre for Evidence Based Medicine, Oxford, UK, 2001.
http://cebm.jr2.ox.ac.uk/docs/levels.html

The review process. Those working in HTA agree that outside review of an assessment report is a very important part of the assessment process. Having such a review is considered to improve the quality of the report, although there is no direct evidence that this is in fact the case. However, there are other important reasons for having a structured review process. Clinicians will often feel that assessment has not taken their perspectives into account. An assessment that is no more than a synthesis of existing scientific literature may not be accepted by clinicians. On the other hand, if clinicians are invited to comment on a draft report, **and their comments are used appropriately**, the clinicians gain a sense of "ownership". Furthermore, those working in HTA feel that

reports are enriched by clinical experience, as long as the conclusions are in accord with the scientific evidence.

Box 17. Recommendation grid and standard recommendation language (based on Third U.S. Preventive Services Task Force)

Quality of evidence	Net benefit			
	Substantial	Moderate	Small	Zero/ Negative
Good	A	B	C	D
Fair	B	B	C	D
Poor	I			
A	... ***strongly recommends*** that clinicians routinely provide [X] to eligible patients. (... found ***good evidence*** that [X] ***improves important health outcomes*** and concludes that ***benefits substantially outweigh harms***.)			
B	... ***recommends*** that clinicians routinely provide [X] to eligible patients. (... found ***at least fair evidence*** that [X] ***improves important health outcomes*** and concludes that ***benefits outweigh harms***.)			
C	... ***makes no recommendation*** for or against routine provision of [X]. (... found ***at least fair evidence*** that [X] ***can improve health outcomes*** but concludes the ***balance of the benefits and harms is too close*** to justify a general recommendation.)			
D	... ***recommends against*** routinely providing [X] to asymptomatic patients. (... found ***at least fair evidence*** that [X] is ***ineffective or that harms outweigh benefits***.)			
I	... concludes that the ***evidence is insufficient*** to recommend for or against routinely providing [X]. (Evidence that [X] is effective is lacking, of poor quality, or conflicting and the balance of benefits and harms cannot be determined.)			

Source: Harris RP et al, 2001.

Box 18. Review process

- Did the report undergo an expert review before publication?
- Who reviewed the report (disciplines)? Were there possible conflict(s) of interest?
- Were the comments from reviewers incorporated into the final report? How?
- How many comments were usable? How many were not usable?

Source: Busse et al, 2001.

An assessment report needs a statement concerning the review process. The statement could be structures along the lines of that shown in Box 18.

Depending on the particularly context, the HTA Body may have one or more advisory groups already involved in the assessment. These advisory groups have already undoubtedly given much useful information and furnished different perspectives during the assessment process. Nonetheless, when a draft report is available, it should be sent for a wider review. If the assessors do not intend to take these reviews seriously, however, the reviews should not be solicited. Such a practice only frustrates well-intentioned clinicians who send their comments in good faith. It takes considerable time and effort to deal with reviews. Such effort does pay off. Typically, an in-depth review

Box 19. Grade of recommendations used for Polish practice guidelines[1]			
Grade of recommendation	**Clarity of risk/benefit**	**Methodological strength of supporting evidence**	**Implications**
1A	Risk/benefit clear	Randomised trials without important limitations	Strong recommendation, can apply to most patients in most circumstances without reservation
1 B	Risk/benefit clear	Randomised trials with important limitations (inconsistent results, methodological flaws*)	Strong recommendations, likely to apply to most patients
1 C+	Risk/benefit clear	No RCTs but RCT results can be unequivocally extrapolated, or overwhelming evidence from observational studies	Strong recommendation, can apply to most patients in most circumstances
1 C	Risk/benefit clear	Observational studies	Intermediate strength recommendation; may change when stronger evidence available
2A	Risk/benefit unclear	Randomised trials without important limitations	Intermediate strength recommendation, best action may differ depending on circumstances or patients' or societal values
2 B	Risk/benefit unclear	Randomised trials with important limitations (inconsistent results, methodological flaws)	Weak recommendation, alternative approaches likely to be better for some patients under some circumstances
2 C	Risk/benefit unclear	Observational studies	Very weak recommendations; other alternatives may be equally reasonable

*These situations include RCTs with both lack of blinding and subjective outcomes, where the risk of bias in measurement of outcomes is high, RCTs with large loss to follow up. NOTE: Since studies in categories B & C are flawed, it is likely that most recommendations in these classes will be level 2. The following considerations will bear on whether the recommendation is Grade 1 or 2: the magnitude and precision of the treatment effect, patient's risk of the target event being prevented, the nature of the benefit, and the magnitude of the risk associated with treatment, variability in patient preferences, variability in regional resource availability and health care delivery practices, and cost considerations. Inevitably, weighing these considerations involves subjective judgement. Source is Gordon Guyatt, MD, Hoger Shunemann MD, Deborah Cook, MD, Roman Jaeschke MD, Stephen Pauker, MD, Heiner Bucher, MD. McMaster University, Hamilton, Ontario: Canada.

process will lengthen the assessment time considerably. If there is great pressure for the outcome, the review process may have to be curtailed.

It is important to document this review process. This documentation will head off comments such as that only part of the literature has been used or that literature from academic medical centres that is not very relevant to community practice was the only source for the assessment. A transparent process of assessment should both seek outside reviews and shown that they have been taken seriously.

Completing the coverage report. The overall coverage report, then, might have these sections:

1. Statement of the policy question
2. Methodology of the coverage report
3. Background information
4. Results
   Safety
   Efficacy/effectiveness
   Psychological/social/ethical considerations
   Organisation/professional considerations
   Economic issues
5. Discussion
6. Conclusions
7. Recommendations

2.2.5 *The Coverage Process*

As previously stated, the actual coverage decision should be taken in an entirely different organisation from that which has carried out the assessment. In this report, these two bodies have been called "the HTA Body" and "the Coverage Body".

The HTA Body is fundamentally a scientifically oriented organisation. It's purpose is to furnish the "facts", the "truth", insofar as possible. In most cases, the HTA Body is part of the insurance system. In Switzerland, for example, the HTA Body is part of the Federal Office of Social Security. HTA's are carried out following a process similar to that described in this report, with the assistance of outside experts, often on contract. In the Netherlands, the HTA Body was earlier part of the Sickness Fund Council, but the organisation and conducting of HTAs was overseen by a specially-appointed committee of experts not otherwise associated with the Council. Perhaps because the policy makers considered that the closeness of the HTA Body to the insurance decisions in the Sickness Funds Council raised the possibility of bias, the HTA Body was moved to a national scientific advisory body in the late 1990s.

This is in accord with a general trend toward independent HTA agencies and programs, to prevent HTA from becoming a purely political function. There is no reason that a national HTA Body could not make these coverage recommendations, although this is a

rather specialised form of HTA. In this project, establishment of a national HTA Body will be recommended for this purpose as well as others.

The Coverage Body, on the other hand, is a political body. It should be in a political setting, or have access to a high level political setting. It should have the power to make coverage decisions. These coverage decisions are enforceable by law. Thus, the Coverage Body is a powerful body. Its structure will be considered later in this report.

In Switzerland, there is a specially appointed Federal Coverage Committee that actually makes the coverage decisions, although the decisions must be accepted by the head of Swiss Social Security. In the Netherlands, the Sickness Funds Council itself makes the coverage decisions, but they must be formally accepted by the Minister of Health.

The figure shows the coverage process schematically. As already indicated, coverage is not automatically made on the basis of efficacy and safety or cost-effectiveness. Coverage is a political process. This is appropriate in a democracy. Important decisions are not made by technocrats or subject-area experts. They are made by people elected by the population to represent their interest, or by those appointed by the elected officials.

Therefore, the most important input to coverage is the policy considerations surrounding a specific technology. In many cases, these considerations may not be very important, but in other cases they will more-or-less totally determine the outcome of the coverage process. From an objective point of view, such considerations as the technological state of the health system of the country may play a large role. From a less objective point of view, the wishes of the population or of administrators and clinicians may play a key role in coverage decisions. As already discussed, the media plays a key role in many decisions. Coverage is a complex process, and the coverage report should be seen as only one input. On the other hand, in most cases, once a process of assessment has been developed that is respected and thought to have integrity has been developed, the report on the technology would be expected to play an important part in most cases.

The Coverage Body must decide exactly what to do about coverage in a particular case, as shown in the figure. There are a series of possibilities.

The first possibility is just to say "NO". Generally, there cannot be a permanent no. New evidence is continually developing. Still, the "no" can also be conditional. The Coverage Body may explicitly say that further assessment might lead to a different decisions. Perhaps industry or others would be interested in further assessment of the technology.

The second possibility is to say "YES", the technology is covered without restrictions or conditions.

The third possibility is to say "YES, CONDITIONALLY. There are a number of reasons to put conditions on a coverage decision. One is to define indications. As mentioned above, a technology may be very beneficial for a certain problem or for a very severe form of that problem, but not for other problems. Therefore, it can be beneficial to furnish coverage only for certain disease indications.

Another condition is to require a certain defined setting. It is known that quality is better with large, technical technology if it is provided at higher volumes because the specialists, the team, and the system gain more experience and develop more skill with higher volumes. So the coverage decision could define certain characteristics that would be required, such as a certain volume of services. Likewise, many modern technologies require special facilities for best quality of services to be provided. For example, cardiac catheterisation is probably best carried out in an institution with a full range of cardiology services. The skills of those providing the service are also important. Commonly, a certain level of training is required in coverage decisions. For example, an endoscopic procedure might only be covered when carried out by an appropriate specialist who has had a special defined course in endoscopy.

Another type of condition is associated with further data collection. It might be that the Coverage Body recognises that the technology shows significant promise, but that the evidence is not yet definitive. The Coverage Body could tie a "yes" decision to a requirement for data collection, such as participating in an RCT or a data bank. In Switzerland, this is a commonly used mechanism with promising technologies where there is significant demand for the technology. All those providing the service are required to submit information, including outcome information, to a databank, which makes reassessment possible later.

As emphasised in other parts of this report, and in the report of Task 1, the coverage process is only one part of attempts to rationalise health care, although it is certainly an important part. An urgent priority is to influence clinician (especially physician) and hospital use of covered technologies. The key point is that coverage is only one step. Methods to influence hospitals and doctors include contracts (including contracts concerning volumes of service), payment regulations, and budgets. They also include directly involving providers in assuring cost-effective care through carrying out HTA and promoting evidence-based medicine. Physicians largely determine the indications for use of technology, so they must be involved in efforts to improve cost-effectiveness and quality of health care. One possibility gaining attention in many countries is clinical guidelines. Guidelines can be developed by many groups, including government, insurance companies and sickness funds, and physician organisations. In all cases, guidelines need to be carefully evaluated before they are used.

2.2.6 *A Simplified Process of Assessment*

The system for making a coverage report described above should fulfil the requirements for a rigorous and transparent process for HTA in relation to coverage decisions. However, the process described here is also complicated and laborious. It is quite similar to that carried out by some organisations, such as BlueCross Blue Shield in the United States (BlueCross Blue Shield, 2000). On the other hand, some organisations have simplified this process considerably. Undoubtedly, Poland will do the same.

A minimum process might be described as follows (Cranovsky et al, 1997):

1. **Identification** of the technology in question. The identification might be done by either the Coverage body or the HTA Body. The Coverage Body would determine if the technology were a potential priority for an assessment.

2. **Literature review** to determine the availability of information. If sufficient information were available, especially in the form of systematic review(s), an assessment report would be developed. If sufficient information were not available, the Coverage Body would make the decision on other grounds, perhaps delaying the decision for months or years.

3. **Synthesis** of available information, including clinical, economic, ethical, and social data. Extensive expert input would be sought in this process. The synthesis leads to judgements and perhaps recommendations. Expert review and criticism would remain an important part of the process.

4. A **coverage proposal** would be developed by the Coverage Body or policy maker based partly on this synthesis. Acceptance of the coverage proposal by policy makers would lead to a coverage decision.

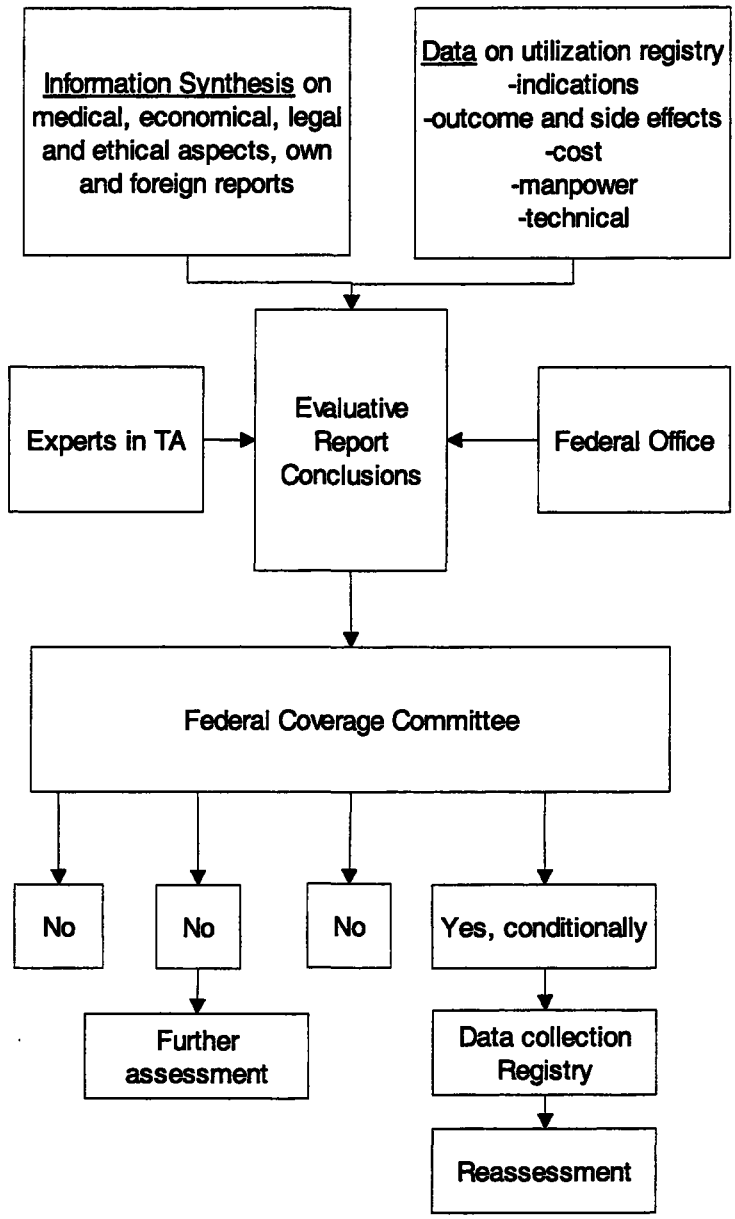
In fact, the Swiss process resembles that is shown in Figure 2 (Cranovsky et al, 1997; Swiss Federal Office of Social Security, 1998). Faced with a demand for coverage of a new technology, the Swiss Federal Office of Social Security will carry out an information synthesis to determine how much information is available on the medical, economic, legal, and ethical aspects of the technology. At the same time, the Insurers Organisation and the Physicians Organisation will be asked if the method is controversial.

The synthesis and the opinions are presented to the Federal Coverage Committee, which makes the decision. Generally speaking, if the technology is not considered controversial by one or both organisations, it will be covered. If it is considered controversial by one or both organisations, the technology will be subject to an assessment. The staff will prepare an HTA based on the literature and other data, as available, including HTA reports from other countries, and after discussion with the applicant, will make a recommendation to the Coverage Committee.

The options of the Coverage Committee in Switzerland are quite similar to those described above in the idealised coverage process. In short, the Swiss system is simplified to a considerable degree, while retaining the essential steps. One omission, however, is the extensive expert input recommended in this report.

In this project, a number of coverage reports have been carried out based on a process similar to this one. This is a demonstration that it is possible to use HTA in coverage decisions without devoting very large resources to the process. The results of these assessments will be presented later in this paper.

Figure 2: Swiss Process of Assessment for Coverage Decisions



Source: Eur-ASSESS Project Subgroup Report on Coverage, reprinted with permission.

2.2.7 *The Coverage Body*

The HTA Body has been described briefly above, and will be described at more length in other reports from the project. But what should be the make-up of the Coverage Body? Coverage decisions could be made by one individual, perhaps in the Ministry of Health, but this would not be desirable, because a transparent process with an open discussion is beneficial in itself.

As emphasised above, the Coverage Body is a political body, that is, it has considerable power and it is subject to many pressures, including the pressures of lobbying by physicians, patients, and industry. The Coverage Body should be appointed by an important policy maker, perhaps the Minister of Health. Formally speaking, it would be

desirable if the Coverage Body made proposals to the Minister of Health, who could accept, modify, or reject them.

The Coverage Body is typically made up of important stakeholders in the health care system, including political figures, insurance and sickness funds representatives, public health experts, clinicians, and administrators. The balance in this committee is obviously important to avoid domination by one group, such as professors. Industry will undoubtedly seek representation in this committee, but to agree seems unwise, since any industry representative will inevitably have a biased position in favour of coverage. The representatives on the Coverage Body should be as objective and free of conflicts of interest as can reasonably be expected.

What about the general public? Eddy (1991) has argued that only consumers (patients and future patients) can best decide which health services are worth their costs. As he says, the public (potential patients) pay for health care through taxes or premiums and they gain the benefits and suffer from the harms. Perhaps the empowerment of an informed consumer is an (at least) partial answer to the problems facing health care policy. Generally speaking, consumers do not have access to good information indicating the degree of ignorance in health care. Furthermore, they are not generally adequately informed on the probabilities of benefits versus the probabilities of being harmed. If patients were told that the experts could not find evidence of benefits, as in the studies of appropriateness described above, would they accept these procedures in uncertain situations? Implementation of a new coverage policy involving respected representatives of the general public offers many advantages.

No recommendation will be made concerning the make-up of the Coverage Body. Such a decision must be acceptable in Poland both administratively and culturally. Such a decision is best made by Polish policy makers.

3 Assessment Conducted by the Polish Association of Quality Promotion Concerning the Basic Benefits Package

3.1 Introduction

Members of the Association of Quality Promotion (TPJ) carried out assessments for the project as part of this Work Package to illustrate the feasibility of developing HTA reports for this purpose on short time-frames and with limited resources. The process was directed by Dr. Krzysztof Landa, who reported to Prof. Assoc. Rafal Nizankowski.

Each expert involved in verifying and assigning technologies to defined categories was obliged to follow methodology developed by TNO and TPJ. Each expert shared responsibility of every assignment done by him or her with another expert who was a reviewer of the assignment. Every assignment document was reviewed by TPJ experts as a crosscheck. The references used are listed in every assignment document. The publications the assignment document was based on are gathered and attached as printed material to the assessments. All completed assessments are submitted with this report as appendices.

3.2 Assumptions

The actual basic benefit package (BBP) was assumed to be all technologies used in Poland at the beginning of the process (currently used technologies).

The project's goals for Working Paper 6 were:

1 To use the present structure of the BBP, and
2 To examine as many technologies as possible by verifying their efficacy, effectiveness, and/or cost-effectiveness.

Because time and human resources available for the realisation of Working Paper 6 were limited, it was known from the beginning that only a small number of all technologies currently used could be verified. Nonetheless, not only did the experience show that the process is feasible, but it produced information that could be used immediately in coverage decisions.

Topic selection was guided by the goals of the project and the limitations mentioned above. Most of the selections were done based on availability of valid secondary studies and some other factors, which had to be balanced against each other. The experts were asked to seek a balance between the following factors:

- technologies used in various medical specialties (selection should have illustrations from as many specialties as possible. A balance toward one or a few medical specialties in the verification and assignment of technologies ought to be avoided),
- diagnostic procedures and treatment technologies,
- technologies directly associated with drug use and no drug technologies,

- number of technologies assigned to individual categories (to give examples of technologies in each category) (see below).

Medical technology is not just a name of a drug, device, method or procedure. Technology needs to be described by population, intervention and outcome. Each technology description requires detailed information on:

- target group - population of patients that technology is applied to (it should be similar to the population of the clinical trials),
- type of intervention, dose, form, method of application, frequency and duration of use etc.,
- aim of the treatment or diagnosis – what the technology is being used for in respect to primary outcome.

3.3 The Structure Used in the Work

The structure was divided into nine categories. The verified technologies were assigned to one of these nine categories. These categories are grouped into two levels: six categories refer to efficacy or effectiveness and three to cost-effectiveness (although if relevant other kinds of economic evaluations may be taken under consideration here) of technologies.

The first level (efficacy or effectiveness) level consists of following categories:

1. effective technology - effectiveness proven in RCT,
2. effective technology – effectiveness known or assumed, not needed to be proven in RCT,
3. effective but risky technology – proven benefit but relatively high (although conditionally acceptable) risk of harm,
4. harmful technology – proven harm by RCT
5. harmful technology – harm not needed to be proven in RCT
6. technology of unknown effectiveness – efficacy or effectiveness studies not done or studies of poor quality (not valid).

An assignment to the first level category refers only to diversification of technologies based on their evaluation in valid clinical trials or lack of such trials. This demonstrates the state of knowledge, what technologies are of proven effectiveness, unproven effectiveness or harmful. The assignments to the first level say nothing about the strength of interventions and do not deal with comparison with other options.

The second level (most often cost-effectiveness or cost minimisation) consists of following categories:

A. relatively highest cost-effectiveness,
B. relatively lower / medium cost-effectiveness,
C. relatively lowest cost-effectiveness.

The category of cost-effectiveness is associated with an obligatory comparison to the most important optional procedures / technologies used to achieve the same primary endpoint. Only technologies from category I, II and III of the first level are taken into consideration as the assignment to the second level matters. Only the assignments to the second level allow one to draw conclusions concerning the compared interventions strength and cost-effectiveness.

Positive list – consists of technologies from categories: I, II and III as well as all technologies from categories: A, B and C of the second level of BBP.

Negative list – consists of technologies from categories: IV, V and VI.

Each assignment to the particular category of the structure consists of a detailed description of technology. Each technology is described by population, intervention, and outcome. Some examples are presented here:

MEDICAL TECHOLOGY								BBP CATEGORY
INTERVENTION				POPULATION			PRIMARY END POINT	
Active treatment	**Dose**	**Route of administration**	**Period of administration**	**Disease**	**ICD -10**	**Age**	**Primary end point**	**BBP category**
Ganglioside	1 x 40 mg	i. m.	28 days	Ischaemic stroke	G46	Adult	Reduction of mortality	VI
Ganglioside	1 x 100 mg	i.v.	15 days	Ischaemic stroke	G46	Adult	Reduction of mortality	VI
Gangliozyd GM1	1 x 100 mg	i. v.	During the first 5 days	Ischaemic stroke	G46	Adult	Reduction of mortality	VI
	1 x 100 mg	i. m.	During the next 25 days					

Active treatment	Dose	Route of administration	Period of administration	Disease	ICD-10	Age	Primary end point	BBP category
Dexamethasone	0,6 mg	i. m.	Single dose	Croup	J 05	Children	Improvement in croup scale Reduction of symptoms	I
Dexamethasone	0,6 mg	p. o.	Single dose	Croup	J 05	Children	Improvement in croup scale Reduction of symptoms	I
Dexamethasone	10 - 20 mg	Nebulisation	Single dose	Croup	J 05	Children	Improvement in croup scale Reduction of symptoms	I
Budesonid	2 mg	Nebulisation	Single dose	Croup	J 05	Children	Improvement in croup scale Reduction of symptoms	I

Active treatment	Number of sessions	Interval between sessions	Volume of plasma exchanged	Disease	ICD-10	Age	Primary end point	BBP category
Plasmapheresis	2	2 days	40 ml/kg	Mild GBS	G 61	Adult	Improvement in disability scale	I
Plasmapheresis	4	2 days	40 ml/kg	Severe GBS	G 61	Adult	Improvement in disability scale	I

Active treatment	Dose	Route of administration	Period of administration	Disease	ICD-10	Age	Primary end point	BBP category
Enoxaparin	2 x 1,0 mg/kg	s. c.	5 – 10 days Until INR > 2,0	Deep vein thrombosis (DVT)	I 80	Adult	Reductions of episodes of recurrent DVT and PE	I
Enoxaparin	2 x 1,5 mg/kg	s. c.	5 – 10 days Until INR > 2,0	Pulmonary embolism (DVT)	I 26	Adult	Reductions of episodes of recurrent DVT and PE	I
Unfractionated heparin	Initial dose 80 UI/kg	i. v.	5 – 10 days Until INR > 2,0	Deep vein thrombosis (DVT)	I 80	Adult	Reductions of episodes of recurrent DVT and PE	II
	18 IU/kg/h	i. v.						

Active treat-ment	Dose	Route of admin istra-tion	Perio d of admin istra-tion	Disease	ICD -10	Type of di-sease	Primary end point	BBP category
Inter-feron beta-1a (Rebif)	22 µg trzy razy w tygodn iu	s.c.	24 month s	Sclerosis multiplex	G 35	Relap sing-remitti ng	Reduction of exacerbation	I
Inter-feron beta-1a (Rebif)	22 µg trzy razy w tygodn iu	s.c.	48 weeks	Sclerosis multiplex	G 35	Relap sing-remitti ng	Reduction of exacerbation	VI
Inter-feron beta-1a (Rebif)	44 µg trzy razy w tygodn iu	s.c.	36 month s	Sclerosis multiplex	G 35	Secon dary progre ssive	Delay of progression	VI
Inter-feron beta-1a (Rebif)	22 µg trzy razy w tygodn iu	s.c.	36 month s	Sclerosis multiplex	G 35	Secon dary progre ssive	Reduction of exacerbation	I

An assignment document was developed for all cases. In most of the cases, the assignment documents structure covers the following:

1. Title page
   - title of the assignment document,
   - number of the assignment which consists of: ICD-10 number / serial number / intervention / "0" if there is lack of cost-effectiveness analysis or "1" if cost-effectiveness analysis is present,
   - the Association and TNO logos,
   - names of the authors and the reviewer,
2. Contents,
3. Table of the assignments (all technologies covered by the assignment document in a way of intervention / population / outcome),
4. Assumptions for the analysis,
5. Detailed description of:
   - medical problem,
   - population,
   - intervention,
   - primary endpoints taken into consideration,
6. Search strategy,
7. List of searched data bases,
8. Criteria of inclusion of clinical trials to the analysis,
9. Effectiveness analysis:
   - efficacy (benefit expected from the intervention)
   - safety (risk / harms associated with the intervention),
10. Conclusions,
11. Bibliography – publications grouped to included and rejected papers.

The structure of some assignment documents is different from the one described above due to specificity of the topic. Some of the chapters are excluded in some assignment documents and some new chapters may be included.

3.4 Process

The process of assessment and reaching conclusions within this project was simplified in respect to the assumptions described above.

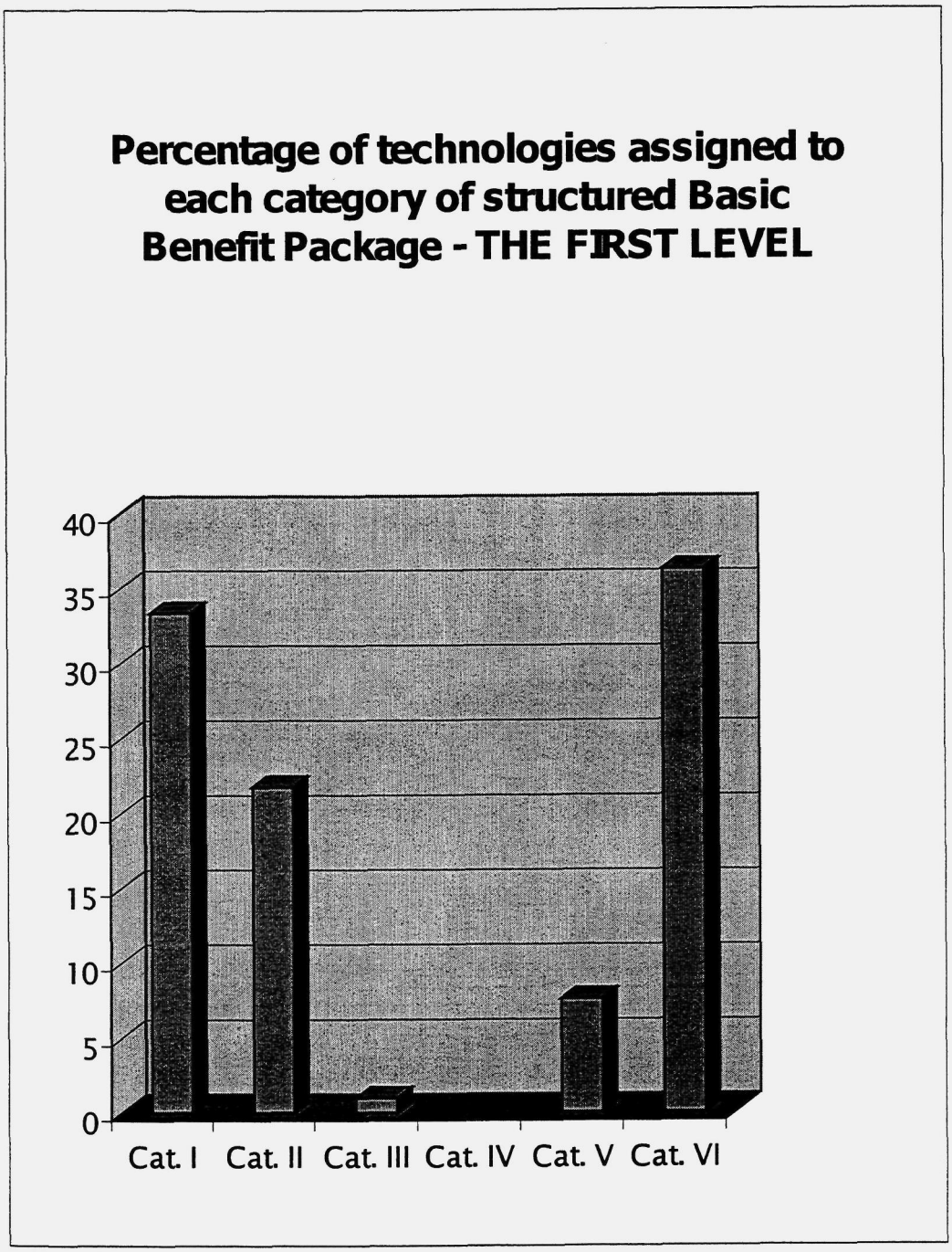
1 Topic selection

Topic selection in respect to the verification of effectiveness or cost-effectiveness of currently used technologies should be the responsibility of the Coverage Body, the HTA Body with the agreement of the Coverage Body, or both of them together. Concerning the use of the benefit package as a tool in reimbursement policy, technologies to be evaluated and assigned to a category should be carefully selected due to their importance to the society. Verification and assignment of currently used technologies is a time consuming and requires involvement of trained specialists in the process. Due to the goals (most of all the introduction of the BBP structure and to achieve a sufficient number and variety of assigned technologies to make possible the proper presentation of the BBP as a tool for reimbursement policy making) and limitations (time and human resources) of the project, topic selection was partially random, to assure keeping the balance between some factors and partially dictated by availability of valid secondary studies.

Topic selection was done by the individual experts of the Association, subject to approval of the supervisors, because of the lack of involvement of a Coverage Body and the lack of a governmental HTA Agency in Poland. In short, topic selection was done differently than it otherwise would be for practical reasons since the objective was to demonstrate the feasibility of the approach. To achieve as great a number of technologies as possible, individual experts relied in most of the cases on the availability of valid literature and expertise. At the second level of the BBP structure, only studies with Polish cost data were taken under consideration.

In the task realisation the Association experts achieved a good balance of the factors described above (factors taken into consideration in the practice of topic selection). However, too few technologies ended up in categories III and IV (see Figure 3). Assignments to categories III and IV will be priority in the continuation of the project.

Figure 3: Category Assignment



2 Literature review / availability of information

The process of literature review largely depended on the available of systematic reviews of the literature, updated by primary literature review. Therefore, the review of literature on primary research was rather limited.

Verification and assignment of technologies were done through the following process:

a) search for secondary studies on effectiveness – assessment of their validity – search for more recent primary studies not included in the secondary studies – compilation or comparison of the results from secondary studies and primary clinical trials – preparation of the assignment document – inclusion to the relevant category of BBP structure (most frequent mode),

b) search for primary studies for technology – assessment of their validity – comparison of the primary studies results – preparation of the assignment document – inclusion of technologies to the relevant category of BBP structure (because of time consuming process this mode was used in just few cases),

c) in respect to the second level categories of cost-effectiveness, all of the assignments were based on the evaluations done prior to this project by the Association.

The most often used secondary studies the assignments were based on:
- systematic reviews of the Cochrane Collaboration,
- effectiveness analyses of Prescrire International (applying mainly to drugs).

About 40 assignments are based on primary research studies. The synthesis was done for the project purposes.

3.5 Human Resources, Timetable and Final Outcome

Twenty members of the Association worked on the assignments. They were not employed full-time in the project. In summary, 11 full time positions (plus the involvement of Prof. Nizankowski and the Logistic Team of the Association) were available for the work on the assessment in the months of October, November and December 2001, and January and February 2002. They were medical doctors, economists, a statistician and a medical physicist; all trained in EBM and HTA.

Some of the comparative cost-effectiveness analyses done by the Association before the beginning of this project were used in the BBP elaboration process. This should be treated as a contribution of the Association to the project.

The experts were able to verify and assign 320 technologies to the first and the second level categories. As noted above, this does not mean that 320 individual names of technologies are given. Instead, the 320 technologies refer to the specific populations to be treated, dosages, etc. Each individually named technology has several of these listings, totalling in all cases to about 450.

4 Basic Benefits Package - Implications for the Future

Parts 1 and 2 of this report have laid out a detailed process for doing health technology assessment as a basis for health insurance benefit coverage decisions. Part 2 has also given a simplified process that could be used. Part 3 has tested the feasibility of using this process in the Polish context. This working paper has shown conclusively that such assessments can be done in Poland. A procedure has been set up in Poland that is usable for this purpose (the details are given in Part 3). Naturally, the procedure, including the definition of categories and assignment to them, might not be exactly as described in this paper. To repeat, the important point is that this paper shows **that it can be done in Poland.**

At the same time, the experience has confirmed statements in Working Paper 1 that to develop a benefit package based on HTA is a long-term endeavour. For eleven full-time people to assessment 320 technological applications in a period of a few months is a great achievement. But it must be remembered that the 320 assessments cover only a small percentage of all health technologies.

The technical implication of this experience is that strategic decisions need to be made if such a system is to be developed in Poland. Since all technologies cannot be assessed in the short-run, priorities would have to be set. What is the basis for such priorities. Some options would include:

1. Assessing all new technologies before they are paid for,
2. Running a formal priority-setting process as sketched in this working paper to determine which technologies should be assessed,
3. Identifying technologies already in place that need to be assessed and setting priorities between them,
4. Assessing technologies that might make a significant contribution to the health of the Polish population. For example, preventive procedures seem seriously under-emphasised now. Prevention could be assessed as an aid to decision-making in this area,
5. A combination of these approaches could be used. In fact, all of the first four methods could be combined, which is the situation in the Netherlands.

To set up such a process and assure the use of HTA in developing a benefit package requires political decisions and legal regulations. The health insurance law may need to be changed if efficacy/effectiveness, as determined by scientific evidence, is to be an important criterion for coverage. Otherwise, regulations will be needed and some new organisations are required to carry out the process.

A HTA Body or agency is needed at the national level in this case. Such a body would not only produce independent, objective, evidence-based reports about the effectiveness and, if possible, the cost-effectiveness of health technologies, but it would also deal with BBP elaboration and updating. Where could such an agency be located administratively? The two main options would be associated with the Ministry of Health or with the Sickness funds. Since the sickness fund association has been abolished and its functions taken over by the Ministry of Health, a location in the Ministry seems the only feasible option at the moment.

The structure and functioning of the HTA Body in general and with regard to the benefit package need to be specified. The procedures must be transparent and stated clearly before the HTA Agency is legally established. The results of assessments for purposes of coverage are of interest to a number of bodies:

1. The Coverage Body – in reimbursement policy (described below).

2. Sickness Funds – they contract only technologies from the positive list, although they may decide not to cover all of them, especially from categories III, B and C.

3. Additional health insurance companies – they may contract not only technologies from positive list contracted by the Sick Funds but also not covered or partially covered by public health insurance.

4. Medical professionals – they refer to the BBP as the source of valid evidence-based information. Medical professionals actively seek for information from BBP but also receive it from the HTA Body and insurance companies interested in changing practice (such information is proven in many countries to have great impact on practice, improvement of quality of care and rationalisation of costs).

5. Patients – they could also look to the BBP as a "guarantee" of services that are a right for everyone. This would give a level of protection of human rights to care that presently does not exist in Poland.

6. Media.

7. Scientists.

To be successful, the HTA Body must closely cooperate with the Coverage Body. The make-up of the Coverage Body is of critical importance. As suggested in this working paper, the Coverage Body would be established by the government and could include a variety of different publics. This body would make the actual decisions, perhaps under the authority of the Minister of Health, about which services should be reimbursed by public health insurance, which should be banned and which services may be contracted within additional health insurance. The Coverage Body would use at its reference points the assignments of technologies to the categories of BBP, which would be managed by the HTA Body.

The two bodies would work according to the following simplified model for a proposed health technology assessment process for making health insurance coverage decisions. This has already been described in Part 1, but is somewhat modified here to take into account the experiences of this project.

- Identification of the technology in question, by the Coverage Body or by the Health Technology Assessment Body. The Coverage Body must determine if the technology is a potential priority for an assessment. If so, it would ask the health technology assessment body for an assessment. The HTA Body could do the assessments on its own or could able to contract with another body for the

assessment (under supervision from the HTA Body). All assessments would be evaluated in respect to their quality due to the clear, objective and transparent criteria. The set of validity criteria should refer to effectiveness and cost-effectiveness analyses.

- Assessments should consist of scientific literature review to determine the availability of information whether the service is proven effective, and, if possible, whether it is cost-effective. If sufficient information for a health technology assessment is available, the health technology assessment body would propose an assessment to the Coverage Body, including a time frame for the assessment. If not, the Health Technology Assessment Body would propose supporting a prospective, well-designed study, such as a randomised clinical trial, and perhaps a simultaneous cost-effectiveness analysis to develop new information. If this proposal is accepted, a delay of several years until the final decision can be expected. Naturally, the Coverage Body may reject the option for an assessment and make the decision on other grounds. It can also, depending on the characteristics of a technology and of the clinical condition to which assessment will be applied, propose temporary conditions with coverage (as in the Netherlands and Switzerland).
- When the assessment and/or systematic review is done the Health Technology Assessment Body would assign technologies to the appropriate categories of the BBP and inform the Coverage Body of the inclusion as well as the outcomes of the evaluation.

The assessment would be based on a synthesis of available information, including that on efficacy, safety, efficiency, and social and ethical aspects, leading to a complete assessment report. Extensive expert input would be sought in this synthesis process. The synthesis leads to judgements, conclusions, and (perhaps) recommendations from the Health Technology Assessment Body or program to the Coverage Body. While the Coverage Body obviously needs clear indications concerning the value of a technology, whether recommendations are appropriate depends on the specific context. The Coverage Body may prefer not to have recommendations, since it has the effect of bringing pressure for a specific action. On the other hand, the Coverage Body may ask for recommendations to consider.

A coverage proposal is then developed by the Coverage Body. After expert review, including review by the HTA Body, the coverage decision is made, published, and promulgated.

Technologies from the positive list (categories I, II, III and A, B, C) should be considered for coverage from public insurance in the first place. Technologies from the positive list not covered by the public health insurance might be contracted and offered by an additional health insurance. Technologies from category C must be carefully considered before coverage decision is made.

Technologies from the negative list of categories IV and V should be banned. Technologies from category VI should be discouraged from use and their reimbursement from public resources ought to be limited.

Such a process would undoubtedly improve Polish health care services in terms of both efficacy and cost-effectiveness. The process would help assure that limited health care resources are well spent.

5 Discussion

The issue of health benefits coverage has gained increasing attention in recent years. As described in Working Paper 1, a number of European countries (as well as countries in other parts of the world) have implemented partial or more complete systems of coverage based on health technology assessment. Economic constraints on health care have forced countries to deal explicitly with costs and cost-effectiveness of health care. HTA offers information that can help in such decisions.

Coverage is an important issue in health policy. Health policy faces contradictory demands and choices at this time, including controlling costs of care, improving quality of care, and extending access to care. Health policy choices must be based increasingly on high quality information on these different aspects of care. HTA can furnish such information. The goal of high quality care at reasonable cost can only be reached if safe and effective interventions are stimulated and the use of ineffective, unsafe, or inefficient interventions are discouraged. Such stimulation and discouragement can be made with the help of HTA in coverage decisions.

A rigorous process of defining coverage is necessary. Without such a rigorous process, progress cannot be made in this area. In such a case, many ineffective technologies will continue in widespread practice, while new technologies that are effective and cost-effective will not be provided. In practice, Poland must limit available services because of limited resources for health care. Legally, within the European Union, such limitation is not possible without clear, transparent, and well-documented justification. Therefore, Poland's future demands a process similar to that laid out in this working paper.

6 References

BlueCross BlueShield Association. TEC Bulletin. 2000; 17: No. 4. Information 1.888.TEC.4321. BlueCross BlueShield Website at http://www.bluecares.com.

Busse R, Velasco M, Perleth M, Orvain J. Best practice in undertaking and reporting HTA. Report from Working 4, ECHTA project. Stockholm: SBU, 2001 (not publicly available until after 15 October 2001.

Cranovsky R, Matillon Y, Banta HD et al. EUR-ASSESS project subgroup on coverage. International Journal of Technology Assessment in Health Care 1997; 13: 287-332.

Eddy D. Benefit language, criteria that will improve quality while reducing costs. Journal of the American Medical Association 1996; 275: 650-657.

Eddy D. What care is 'essential'? What services are 'basic'? Journal of the American Medical Association 1991; 265: 782-788.

Harris R, Helfand M, Woolf SH et al. Current methods of the U.S. Preventive Services Task Force: A review of the process. American Journal of Preventive Medicine 2001;20:21-35.

Ministry of Health and Social Welfare. Terms of reference: Standard of health care services purchased in the national health insurance system. Warsaw, Poland, 1999.

Office of Technology Assessment. Policy implications of the computed tomography (CT) scanner. Washington DC: US Government Printing Office, 1978.

Petitti D. Meta-analysis, decision analysis and cost-effectiveness analysis. Oxford: Oxford University Press, 1994.

Swiss Federal Office of Social Security. Manual for the standardisation of clinical and economic evaluation of medical technology. Draft 1998. Berne.

TNO Prevention and Health. Standard of health services purposed in the national health insurance system. Leiden, 2000.

A Websites Related to Health Technology Assessment

Bibliographic data bases

INTERNET GRATEFUL MED
http://igm.nlm.nih.gov/

PUBMED
http://www.ncbi.nlm.nih.gov/Pub-Med/

Guides to evidence

ScHARR-Lock's Guide to the Evidence

http://www.shef.ac.uk/uni/academic/R-Z/scharr/ir/scebm.html

Evidence based medicine

COCHRANE COLLABORATION
http://hiru.mcmaster.ca/cochrane/default.htm

COCHRANE LIBRARY
http://www.update-softward.com/ccweb/cochrane/cdsr/htm

CENTRE FOR EVIDENCE BASED MEDICINE
http://cebm.jr2.ox.ac.uk/

YORK CENTRE FOR REVIEWS AND DISSEMINATION
http://www.york.ac.uk/inst/crd/

NETTING THE EVIDENCE
http://www.shef.ac.uk/uni/academic/R-Z/scharr/ir/netting.html

BANDOLIER
http://www.bmj.com/index.shtml

EVIDENCE BASED PURCHASING
http://www.epi.bris.ac.uk/rd/publicat/ebpurch.index.htm

Health technology assessment

International Network of Agencies for Health Technology Assessment (INAHTA) (provides access to HTA agencies that belong to INAHTA)
http://www.inahta.org

DARE database of abstracts of reviews of effectiveness
http://www.agatha.york.ac.uk

NHS Centre for Reviews and Dissemination (HTA database available here)
http://www.york.ac.uk/inst/crd

Polish Society of Quality Assessment
http://www.cmj.org.pl

TRIP database (allows searching in EBM-related databases,
http://www.tripdatabase.com

Clinical guidelines
http://www.guidelines.gov