Health Technology Assessment (HTA) Guideline

Indonesian Health Technology
Assessment Committee (InaHTAC)
Ministry of Health
Republic of Indonesia

Jakarta, 2017

Health Technology Assessment (HTA) Guideline

Indonesian Health Technology Assessment Committee (InaHTAC)

Ministry of Health
Republic of Indonesia

Jakarta, 2017

Health Technology Assessment

Guideline

Indonesian HTA Committee

Ministry of Health
Republic of Indonesia

Jakarta, 2017

Remarks

National Health Insurance (*Jaminan Kesehatan Nasional*, hereafter referred to as JKN) implementation is based on several laws, including Law Number 40 year 2004 with regards to the National Social Security System, Law Number 24 year 2011 with regards to the Social Security Administrative Body, and others. JKN oversees the quality and financing of healthcare, with the aim of providing access to health care and financial protection for all Indonesians.

The Presidential Decree Number 12 year 2013 regarding Nation-al Health Insurance stated that the development of health tech-nology in JKN should be adjusted to meet medical needs, based on the health technology assessment (HTA) results. Therefore, the Minister of Health established the Indonesian HTA Commit-tee. This Committee is responsible for providing policy recom-mendations to the Minister.
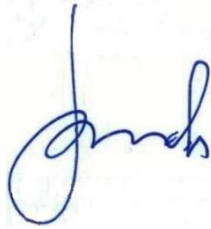
HTA is a comprehensive report on safety, efficacy, effectiveness, economic analysis, sociocultural values, and religion (if needed), as well as budget impact analysis, with regards to health care in this country. The main challenges in this report were evaluation of costs, determining the current benefit packages, and ensuring that health technologies used in JKN are based on evidence (evidence-based medicine) with appropriate financing. In the future, policy recommendations based on HTA results may relate to procurement and funding of health technologies which are cost-effective, or discontinuing the use of health technologies which are not cost-effective.

This book aims to provide reference, guidance, and reading material for all stakeholders involved in JKN implementation.

Many thanks to all parties who contributed to the preparation of this book. May we continue our best efforts for the Indonesian people and their right to accessible health care.

Jakarta, April 2017

Head of Center for Health Financing and Health Insurance

dr. Kalsum Komaryani, MPPM

Preface

Indonesia launched the National Health Insurance (*Jaminan Kesehatan Nasional* / JKN) program on January 1st, 2014, with the goal of providing universal health coverage (UHC) to all Indonesians by 2019. The JKN program is the enactment of a social security system that aims to fulfill basic life needs, e.g., the right to health and protection for every citizen.

JKN is the realization of a government commitment to implement the mandate of the 1945 Constitution (UUD 1945) article 134, paragraph 2, Law no. 40, year 2004 on a National Social Security System. The implementation of UHC or health insurance for all in many countries has increased the level of safety and productivity of their citizens. However, the rapid development of medical technol-ogy, which is increasingly costly, will eventually require an enor-mous budget, while the available funds are quite limited.

The Presidential Regulation no. 12, year 2013 on Health Insurance states that in order to guarantee quality and cost control, the Minis-ter is responsible for conducting a health technology assessment (HTA). Based on the HTA findings, the use of technology for health insurance benefits should be adjusted according to patients' medical needs. As such, the Minister of Health established the Indonesian Health Technology Assessment Committee (InaHTAC) through the Minister of Health Decree no. 171/Menkes/SK/ IV/2014 and no. HK.02.02/ MENKES/422/2016.

The InaHTAC is responsible for providing health policy recommen-dations to the Minister, regarding which health technologies will be covered in the benefit package or become prioritized in the JKN program. HTA is conducted to objectively ensure that JKN benefi-ciaries, i.e., Indonesian citizens, receive the health technology they require. The InaHTAC is tasked with considering the magnitude of health technology needs, the safety and effectiveness of the technologies, and the ability of JKN to fund such technologies.

It was beyond our expectation that only a few weeks after the establishment of InaHTAC, a number of international organizations started to contact us to offer collaboration and support. With the support from those organizations, i.e., WHO (World Health Organ-ization), AIPHSS (the Australia-Indonesia Partnership for Health Systems Strengthening), HITAP (the Health Intervention and Tech-nology Assessment Program), NICE (the National Institute of Health and Care Excellence), iDSI (the International Decision Sup-port Initiative), PATH (the Program for Appropriate Technology in Health (PATH), ADP (the Access and Delivery Partnership), the InaHTAC was able to start working properly.

We understand that in the future it is impossible that all HTA can be conducted by the InaHTAC. There must be other institutions or groups that conduct the assessment using the same principles and process. For that reason, we have developed this HTA guideline to assist in the implementation of HTA for all stakeholders.

This guideline consists of two parts. Part One is a general description of HTA with specific attention to clinical aspects. Part Two provides guidelines for economic evaluation. A third part, which describes step-by-step guidelines to carry out HTA in Indonesia, will be published separately from this document, together with the Minister's Decree on InaHTAC.

This guideline will require further revision in the future. Input from the readers and stakeholders will be taken into consideration for the next edition.

Sudigdo Sastroasmoro
Mardiati Nadjib

Acknowledgments

Table of Contents

Part One

General Description of Health Technology Assessment with Special Reference to Clinical Aspects

Chapter 1
Definition and History of Health Technology Assessment

Definition of health technology assessment

Health technology assessment or **HTA** is increasingly popular among medical and health fields. In general, HTA refers to any kind of effort to improve the quality of health services, including promotion, prevention, diagnosis, treatment, rehabilitation, and long-term care. With the progressive efforts of universal health coverage (UHC) as suggested by the WHO, HTA has become compulsory in every country, while it was merely a suggestion a few decades ago.

There are many definitions of **technology, health technology,** and **health technology assessment** that we can obtain in the literature, but these definitions vary widely. It seems that every institution tends to create its own definitions, even though the core definition is similar. In comparing many definitions, we concluded that some experts assume HTA to be a research activity, while several consider HTA to be methods, and the rest argue that HTA is merely a process. Herein, we summarize the definitions of technology, health technology, and health technology assessment in a brief, complete, and contextually suitable with the current conditions:

- In general, **technology** is defined as the application of scientific knowledge for practical purposes

- **Health technology** is described as all types of interventions used in the health field for promotion, prevention, screening, diag-nosis, treatment, rehabilitation, and long-term care. Health tech-nology includes drugs, biological substances, medical/surgical procedures, support systems, organizations, and management.

- **Health technology assessment (HTA)** refers to systematic evaluation of properties, effects, and the impact of diffusion and use of health technology. It is a multidisciplinary process to evaluate safety, efficacy, effectiveness, as well as social, economic, organizational, legal, ethical, cultural, and religious parameters.

Based on the above definitions and others as mentioned in the references, we concluded that the word "technology" does not only represent medical devices, such as ultrasonography (USG), magnetic resonance imaging (MRI), or positron emission tomography (PET). Health technology includes all types of procedures used in various aspects of the medical field, from promotion to rehabilitation.

The role of HTA in improving the quality of health services

Health technology is rapidly evolving over time. People continue to search for ways to improve health service quality. This drive may be due to dissatisfaction with the current system. In the era of evidence-based medicine (EBM), a step-by-step approach is used to deal with any problems in the medical or health field (for instance, lack of awareness of the negative impact of smoking, the importance of exercise, high maternal mortality rate, difficulties in diagnosing certain diseases, ineffective treatment, etc.).

- The first group to be relied on is researchers. They can offer the options by providing scientific evidence based on their research to reduce the magnitude of the problem.

- Research is a costly undertaking, requiring special facilities and experts who may not be available in daily practice. Alternatively, small scale research may not be useful for detecting potentially high risk adverse effects which happen rarely. In addition, alternative solutions may not differ much from the existing standard approaches. In order to assess these various aspects, an HTA is required for a comprehensive and systematical review.

- The results, after being adjusted to fit with local conditions, could be used to develop or revise clinical practice guidelines in health facilities or hospitals. Clinical practice guidelines with appropri-ate disclaimers ought to be followed by professionals/health providers after the guidelines have been inaugurated by the heads of facilities.

- Eventually, another process is required to confirm whether the health providers followed through with their tasks, i.e., by a clin-ical audit process.

It is obvious that HTA occupies a very important place in evidence-based health care that is patient oriented and focuses on all aspects of distribution and implementation of health technology.

Value, not merely service quality

The development of advanced health services has strengthened other service dimensions, the so-called value of health services, in community as well as in clinical aspects. The value of a service should be in line with its quality, but not necessarily defined by its cost.

Quality improvement is characterized by decreased mortality and morbid-ity, increased quality of life, client/patient satisfaction, and the increased community health status. As such, cost does not only mean monetary pricing, but also the availability of facilities, human resources, time, etc. A premium health service with high cost, like in the US, has a lower value than health service of similar quality but lower cost, such as in North Europe and some developing countries.

In planning for UHC, HTA plays role in utilizing high quality health tech-nology while limiting the demand on resources, thereby allowing every citizen to acquire equal health services. Clinical assessment (especially safety and effectiveness) and economic evaluation (from the modest to the most complicated) are the most frequent aspects to be evaluated in HTA, as will be discussed in the next chapters.

Why HTA is necessary?

There are numerous reasons for the necessity of ongoing HTA. Since every country faces different situations, HTA in one country may not be appli-cable to another country. Likewise, HTA is not valid indefinitely; revision is necessary depending on the nature of the topic.

A few reasons on why HTA should be done:

- Rapid development of medical science and technology, including the development of specializations and sub-specializations, each of which have distinct characteristics and requirements.

- Limited economic resources, not only in developing, but also in developed countries. The application of certain technology always has an economic dimension; the more complicated and sophisticated the technology, the higher the cost.

- Abundant evidence shows that people still use old health technology that is no longer useful or may even be harmful.

- Other evidence shows that many technologies are useful, but are not used, underused, or put to use very late in the utilization of health services.

These reasons can be explained in more detail or added to. For example, HTA is important for the development of universal health coverage, health insurance, changes in the epidemiology of diseases (e.g., increased aging population), and the increase of malpractice claims. Competition among hospitals, especially private ones, also require HTA. Hospitals want to give an impression of offering sophisticated services, which may lead to a tendency to perform unnecessary examinations and/or treatments. The increase in community awareness to demand a quality but affordable health service, is also a reason for HTA.

The main issue is the conflict between the need for technology application and the cost. Prioritization is necessary to decide which technology should

be used, which should be limited, and which should be abandoned. Only an objective and independent assessment can answer the challenge of applying a health technology in the most useful and cost-efficient manner.

History of HTA

The logical evaluation of science has been implemented in patient service since helpers of the sick existed. **Imhotep** (27th century BC) of Egypt, the first known physician in history, treated patients using his knowledge (at that time) and evaluated treatment success and negative effects of medication. A similar analogy can be applied in every generation of doctors through the ages, even though treatments were evaluated in unsystematic observations because research methods, especially clinical trials, were unrecognized.

James Lind, in the mid-18th century, administered 6 different treatments to patients with scurvy; he assessed the treatment outcome, any adverse events, and aspects of availability and drug price. Therefore, in all eras of medical history, evaluation of health technology has been and always been done, according to the level of medical science at the time.

However, a structured, systematic HTA has only been identified since 1960s, after the use of health technology was recognized for its positive and negative effects, as well as need for resources, including financing. The term "technology assessment" appeared in 1965 when it was first discussed at the US Congress. The topic then evolved until 1973, leading the Office of Technology Assessment (OTA) to commence its medical activities in 1975.

At the beginning, HTA studied the safety, effectiveness, cost, and other issues related to the application of health technology. The OTA issued a bioequivalence report in 1974 and its official report in 1976. Since then, HTA has also included social, ethical, legal, and political aspects. The topics assessed also included contraception, organ transplantation, artificial or-gans, ventilator use, genetics, and stem cells. However, though it has involved many things, few HTAs provide a comprehensive report, as only certain aspects were included.

With regards to institutional development, the main task and function of HTA has developed through several stages. First, **the** main task and function of the **old HTA** was to assess whether certain drugs or procedures were safe and effective for clients. In 1990, **the New HTA** was developed, with its main task and function being to answer the following question: "Is this technology cost-effective for improving quality of life or decreasing the mortality rate of the beneficiaries?" Next, **the need-based HTA** assigned a task to assess the impact of technology on people in need (not only individuals) of a certain intervention. Eventually in 1999, the concept of evidence-based medicine arose and a new approach developed: **the Evi-dence-Based Health Technology Assessment,** which summarizes the previous stages and analyzes using an EBM approach.

In Indonesia, HTA was first established as part of the policy analysis unit in the Directorate General of Medical Services, Ministry of Health in 2003, and named the **Technical Team of Health Technology Assessment**. Dis-cussion about HTA had been ongoing since 2001, when a seminar was held in celebration of the 40th anniversary of Fatmawati Hospital with the theme: **Health Technology Assessment, Evidence-Based Medicine, and Clinical Governance.**

In 2002, by the invitation from the Head of Division of HTA and Clinical Practice Guidelines, Ministry of Health – Malaysia, a few Indonesian doctors underwent HTA training in Johor Bahru. After the training, the Technical Team of HTA was established by the Ministry of Health. Although their reports went unpublished up to 2013, the HTA technical team assessed more than 33 health technologies (unfortunately, without formal economic evaluation). Several of these were adopted into national policies, such as HTA of vitamin K injection for newborns. The problem was that the results were not disseminated properly due to a limited budget. Officially, HTA Indonesia has been accepted as a member of HTA International and has been involved in international forums.

Chapter 2

Classification and Scope of HTA

Classification of health technology

One might think that the term **technology** is closely related to technical equipment. However, as explained in the definition, the term technology in the context of health services involves wider concepts. However, all health technology does not need to be assessed in all situations.

Health technology may be classified based on several parameters which were detailed in the health technology and HTA definitions mentioned in **Chapter 1**. Health technology may be classified based on:

- Type of technology

- Purpose or use of technology

- Development and application of technology.

Classification based on the type of technology

- Drugs, such as antibiotics, aspirin, or statins

- Biological matter, such as vaccines, blood products, or stem cells

- Devices, such as pacemakers, or diagnostic kits

- Medical and surgical procedures

- Support systems, such as electronic medical record systems, tele-medicine, drug formularies, or blood banks

- Organizational and managerial systems, such as insurance, or diagnostic related group (DRG)

Classification based on purpose, function or application

- **Promotive**: health activities which prioritize health awareness, promotion of healthy lifestyle, etc.

- **Preventive**: activities which aim to prevent or decrease the risk of disease, or limit the sequelae; e.g., immunization, hospital infection control programs, or fluoride in the water supply

- **Screening**: early detection procedure on patients without any signs/symptoms; e.g., pap smear, mammography, or tuberculin test

- **Diagnostic**: process to determine a disease or medical condition in a subject with clinical signs/symptoms, e.g., electrocardiography (ECG), MRI, or heart catheterization

- **Curative**: treatment with the aim of reducing the signs and /or symptoms, controlling disease, or slowing disease progression

- **Rehabilitation**: activity to restore, maintain, or increase physical or mental capacity of former patients in order to increase functioning, e.g., training program for post-stroke patients, exer-cise for post-heart-attack patients

- **Palliative care**: care which aims to increase the quality of life of the patient facing a threatening illness and that of his family, through reducing and preventing suffering, early detection, pain management, and comprehensive assessment of other problems (physical, psychological, or spiritual).

Classification based on maturity and distribution

- Future technology: still in concept, anticipating future use, or still at a premature stage of development

- Technology in an experimental stage, in animal or model trials.

- Technology in the evaluation stage: application for patient use in certain conditions

- Evidence-based technology: used by service providers in disease management or certain health conditions

- Ancient or underdeveloped: the technology has been replaced, proven to be ineffective, or even harmful.

The scope of HTA

According to its definition and purpose, HTA includes the assessment of a wide spectrum of health technology applications including:

- Technical characteristics

- Safety

- Efficacy

- Effectiveness

- Economic aspects, i.e., value for money and budget impact

- Social, ethico-legal, political, and religious aspects

Technical characteristics

Technical characteristics, especially of medical devices for diagnostic purposes (e.g., MRI, CT scan, hybrid angiocardiography) as well as for therapeutics (e.g., stent, hearing aid, device to close congenital heart defects) need attention, though on-the-spot assessment is rarely held. The assumption is that the producer of medical devices and hospitals/doctors who would use the device would want it to be well-functioning and properly maintained. For medical devices in small hospitals, especially those located in remote areas, it is deemed necessary to consider the specifications, indications, treatment, and that calibration have been done properly. Calibration of diagnostic tools should be monitored by all relevant parties.

Efficacy

Efficacy of a drug or procedure is best assessed by randomized controlled trial (RCT) with highly selected participants so that their characteristics are homogeneous. For instance, in order to assess the impact of a new anti-diabetic drug, the selected participants should be diabetic patients with no hypertension, normal cholesterol level, normal kidney and liver function, etc. Furthermore, the participants should be treated in idealized circum-stances. This kind of clinical trial provides strong evidence on the rela-tionship between the drugs and the outcome (blood glucose level) in ideal conditions; i.e., the trial has **good internal validity**. However, this type of trial does not have **good external validity** because in daily practice, diabetic patients often have other accompanying diseases. Therefore, the results of efficacy trials cannot be directly applied in reality. Apart from showing that drug A is more effective than drug B, this type of clinical trial also intends to explain the mechanism process, therefore, this type of clinical trial is also called an **explanatory trial.**

Effectiveness

The effectiveness of drugs or medical or surgical procedures is also best assessed with randomized clinical trials. In contrast to an efficacy trial, clinical trials assessing the effectiveness of certain drugs or procedures do not have strict criteria for subject recruitment. As such, participants are patients similar to those seen in daily practice. The external validity of this type of study is quite good, therefore, the results can be applied in daily practice. In effectiveness clinical trials, the investigators aim to show that drug A is more effective than drug B without trying to explain a mechanism of action, therefore, it is also known as a **pragmatic trial**.

It should be remembered that not all evidence of clinical effectiveness can be or should be obtained through clinical trials. For example, for many decades case series from all over the world showed that without proper surgical treatment, the one-year survival of infants with congenital heart

disease transposition of the great arteries (TGA) was only about 10%. With the development of surgical techniques, anesthesia, intensive care, etc. the one-year survival of such patients increased to 90%. With such a big effect size (80%), clinical trials were unnecessary (or even impossible, not allowed, or unethical). Some experts call this phenomenon an **all-or-none phenomenon**. It is valid and strong evidence for the therapy.

The effectiveness of medical devices, for diagnostic purposes, therapy, or even disease monitoring, is rarely obtained from clinical trials. For instance, the effectiveness of ultrasonography to confirm the diagnosis of subarach-noid bleeding in infants, or of a special device to close persistent ductus arteriosus, atrial septal defect, and ventricular septal defect in infants or children, are usually obtained from case series, not randomized clinical trials. In certain cases, diagnostic research could provide accurate evidence on certain devices or diagnostic procedures by comparing them to the **gold standard**.

Safety

The safety of medical technologies (drugs, devices, procedures, etc.) can be obtained through direct observations, routine reports from hospitals, case reports in the literature, or side effects reports from clinical trials. It should be noted that clinical trials usually involve a few hundred or even less than 100 participants, therefore, rare adverse events (e.g., 1 in 2,000 patients) which are potentially fatal, are not recorded in most clinical trials. With a meta-analysis, since the number of participants is large, then rare side effects may be recorded.

The rare, but potentially fatal, adverse events are frequently found in phase IV clinical trials (post-marketing trials, which is in fact surveillance, instead of clinical trials). Therefore, drugs in clinical trials (phases 1, 2, and 3) that were reported to be effective and safe, hence, allowed to be marketed, a few months or years later, after having been used by hundreds of thousands of patients, may be withdrawn because of rare but potentially fatal adverse events.

It should be remembered that safety is not only assessed for patients, but also for service providers, and the environment. Special attention is required for devices with X-rays and nuclear material. The application of those technologies should meet 100% of the requirements, determined by both the producers and the authorities.

Economic aspect

The application of health technology impacts, with wide variation, economic aspects, in micro- and macro-economics. In **microeconomics**, the cost, price, and payment are related to certain technology applications. For instance, the mean total expenditure per year can be calculated for one thalassemia patient who requires regular blood transfusions; this expenditure can be assessed through a **cost-of-illness analysis**. An economic analysis can be used to compare the cost of two or more technologies with their outcomes, such as **cost minimization, cost benefit, cost effectiveness,** and **cost utility analyses.**

The impact in macroeconomics, as it relates to the application of health technology, includes national expense, resource allocation for health programs and other sectors, impact on insurance companies, investment in health, competition in health services, technology transfer, and employment. The details about economic analysis in HTA will be discussed in Part Two of this book.

Impact on social, legal, ethical, political, and religious aspects

The use of technology clearly impacts other aspects of people's lives, including social, legal, ethical, and even religious aspects. The latter is rarely mentioned in Western literature. However, for Indonesians, religion deserves special attention, since the majority are Muslims. For example, if the production of a vaccine uses lard, then the vaccine would be rejected by the majority. The development of a technology might also pose problems

such as organ transplantation using a living donor or cadaver, breast milk banks, sperm banks, pregnancy termination if the fetus has a confirmed severe congenital defect, and others. As such, these technologies should be assessed properly.

Some social, ethical, and legal aspects can be resolved by local and interna-tional literature searches, however, all should be interpreted based on the social culture, law, and ethics of the local people. Experts should be involved, including religious leaders. Consensus is usually required for decision-making.

Purposes of HTA

There are numerous HTA stakeholders, as all citizens should be concerned with HTA. Usually, the most concerned parties are the policy makers. These are the groups with an interest in HTA:

- The Ministry of Health could apply the HTA results, mainly in relation to National Health Insurance, to determine inclusion of a certain technology into the UHC benefit package.

- The government agencies, such as the National Agency for Food and Drug Control (BPOM), require input on whether a certain health technology can be used (drugs, medical devices).

- The payer (insurance), including the Social Security Health Agency (BPJS), should receive input as to whether certain proce-dures, screening tools, drugs, or devices should be included in the items guaranteed by the insurance.

- Professional health service providers (doctors, dentists, pharma-cists, nurses, midwives, and others) require HTA to obtain valid evidence of whether a certain technology can be used for service.

- Professional organizations can use HTA results to develop or revise the National Guidelines for Medical Practice (PNPK) or Clinical Practice Guidelines (PPK).

- Educational institutions, such as medical faculty / dentistry / public health, and other medical institutions can apply the assessment in the educational process.

- Hospitals, service networks, medical drugs/device providers

- Producers / industries for pharmacies and medical devices

- Parliament or political leaders can use the results for technology innovation, research/development, regulation, insurance, etc.

- Patients whom the service targets are in fact the most concerned. No one wants to be sick or have a serious illness, but when they are, they desire high quality and affordable treatment.

HTA does not necessarily start with its technological aspects, but may have different orientations:

- **Technology-oriented:** for instance, the government would like to assess clinical, economic, or social impact, on professionals or health technology industries (e.g., cancer screening, cochlear implant, other interventions)

- **Problem-oriented:** for example, the development of clinical practice guidelines for certain diseases or medical conditions that require a combination of physical examination and tools.

- **Project-oriented**: for example, to determine whether it is appropriate to purchase a certain device, such as MRI, PET, etc.

These three orientations may stand independently, but may be interrelated.

When to conduct HTA

When should HTA be done? HTA is done not only for new technology, but also the existing technology. There is no standard on when HTA is to be done. In general, the sooner HTA is done, the better the impact. The selec-tion of the topic to be assessed involves many considerations as discussed in **Chapter 3.**

Chapter 3
Methods in HTA

Health technology assessment (HTA) can be done through **primary data** collection or a comprehensive review of **secondary data.** However, a **mix** of both methods is usually the best choice for most HTAs.

Primary data

Primary data means self-collected data (required for HTA) by the HTA team. This includes on-the-spot observation at the service location, in order to confirm whether the device functions well, is safe, etc. However, these spot checks rarely happen. Primary data can be collected from many sources (e.g., hospitals or other health facilities), or by formal research using various possible designs. Theoretically, all types of research, observational as well as experimental, individual or collective, can be used for primary data collection. Research includes cross-sectional studies (descriptive and analytical), case-control studies, cohorts (prospective and retrospective), and even clinical trials, all of which can be done individually or in collabo-ration with other parties, such as universities or research institutions.

Nevertheless, collecting research-based primary data as the main evidence for HTA is not generally done, especially clinical trials. Clinical trials require enormous funding, long duration, and complicated processes. Moreover, definitive evidence may not be obtained with only one clinical trial. Primary data from medical records are frequently used as background research on why HTA (hospital-based or regional) should be conducted.

Primary data for economic analysis must be obtained locally/nationally. Prices for drugs/devices, cost for surgery, doctors' fees, etc., are definitely different among countries and impossible to obtain from the literature, and should be attained from local data.

Secondary data

In many places, HTA uses **secondary data,** also known as an integrative or synthesis method, i.e., a summary of all the existing information or resources. **Integrative literature** consists of literature reviews, systematic reviews, and meta-analyses.

- **Literature reviews.** A traditional literature review is often done unsystematically, lacking a systematic approach in collecting data from literature. In addition, critical appraisal may not be done properly, and the conclusion is frequently not quantitative.

- **Systematic reviews (SR).** Using this method, a literature review can be done systematically. The process begins with systematic and transparent literature searching, with all relevant articles critically appraised, and the results integrated systematically. Therefore, a more definitive conclusion can be achieved. In this SR, a formal statistical analysis was not done.

- **Meta-analysis.** Meta-analysis can be viewed as a systematic review with additional formal statistical analysis. This type is done mainly for clinical trials with quantitative value, but can also be done for many observational and qualitative studies.

For HTA purposes, literature review (unsystematic) should be avoided; it is highly recommended to use SRs and meta-analyses. The question is: how close or similar should the SR and meta-analysis be with the criteria used in the Cochrane database? Developing a SR and/or meta-analysis with prime quality and validity, like a **Cochrane Systematic Review,** needs a long period of time, such as months or even 1 or 2 years to complete the study. Obviously, this is unnecessary for all HTA studies. Imagine if in one year, 20 HTA studies need to be done. Then a proper meta-analysis according to publication standard would not be achieved. Therefore, even though all steps in conducting SR and meta-analysis should be followed, they may not be as rigid as those in SR and meta-analysis for scientific publication purposes.

In 2012, a survey was conducted to know what methods were used to obtain data in 16 HTA committees in Canada, Europe, the US, Latin America, and Australia. The majority of methods used were SR, meta-analysis, and economic modeling. Components that were frequently assessed were effectiveness (more often than efficacy), cost-effectiveness, safety, and quality of life. It also showed variability between countries about the importance of the topics, how to evaluate topics' components, and other aspects, such as the economic assessment guideline, what to do if there is no evidence, and data availability about the developed technology.

Step-by-step of Evidence-Based HTA

HTA was formally started before the introduction of evidence-based medicine (EBM). Since the introduction of EBM in early 1990s, HTA should be conducted accordingly, and known as evidence-based health technology assessment (EB-HTA). To address problems in clinical practice with an EBM approach, then we should do the following:

- Formulate the problem into specific and answerable questions containing the 4 elements of the PICO format:

    a. *Patient* or *Population* or *Problem*;
    b. *Intervention* or *Indicator* or *Index*;
    c. *Comparison*, and
    d. *Outcome* or expected results.

- Use the keywords in the clinical questions to search for valid scientific evidence on the internet;

- Critically appraise the scientific evidence with regards to 3 aspects: validity, importance, and applicability (VIA);

- If the results are good, then the evidence can be applied or recom-mended.

Referring to this general paradigm, a step-by-step approach to evidence-based HTA can be developed as follows:

1   Topic identification and selection

2   Formulating problems into specific questions using the PICO format

3   Determining the methods to be used: integrative method (without economic analysis) or requiring primary data

4   Collecting the necessary primary data

5   Literature search of the internet

6   Critical appraisal of the evidence

7   Synthesis of the appraisal results and recommendations

8   Conclusions and recommendations

9   Disseminating the HTA results

10 Monitoring the recommendations

1. Topic identification and selection

Topics to be assessed in HTA include a wide range of health fields, i.e., promotion, prevention, diagnosis, treatment, rehabilitation, and long-term supportive care.

Frequently, one assessment involves more than one aspect. Any HTA stakeholders may suggest or propose topics, including the HTA Committee itself, Directorates of the Ministry of Health, professional organizations, academicians, hospital associations, insurance (including BPJS), industries (drug/medical device companies), non-governmental organizations, pa-tient associations, and even individuals.

The HTA Committee or any organization which will conduct the assessment, may passively wait for proposals in order to identify topics; however, at the beginning the organization may need to actively reach out to all stakeholders for topic suggestions, after explaining the role of HTA. Experience from other countries has shown that many stakeholders pro-

pose topics within only a few months after the establishment of a formal HTA institution. If the number of suggested topics is more than the capacity, then prioritization is needed.

In general, topics are selected if they meet one or more of the following criteria: (1) high number of cases (high volume), (2) high risk, (3) high cost, and (4) high variability in daily clinical practice. Even if the topic meets only two or even one criterion, if it is deemed important, then an assessment can be done. The following points should be taken into consideration:

- Is the topic necessary for policy implementation? Policies on the inclusion of a drug/medical device in the BPJS benefit package deserve an important place.

- Can the type and number of HTA questions be answered properly, with regards to time and other technical factors?

- Are there sufficient materials or literature to answer the HTA questions?

- Can the results improve health outcomes?

- Do the results have the potential to decrease health service cost?

- Will the results potentially provide input on the social, economic, ethical, political, or religious aspects?

In order to establish a prioritization system for topics, a matrix can be developed, i.e., giving scores to each topic based on the case volume, risk, cost, variability in practice, importance of the results for developing policy, and other relevant aspects, as mentioned above.

2. Formulating HTA questions

HTA questions, also known as research questions, should be developed based on problems that are suggested during topic selection. General questions can be made more specific using the PICO format, as in EBM practice. PICO includes 4 elements: P (Patient or Population or Problem), I (Intervention or Index or Indicator), C (Comparison), and O (Outcome).

For example, a group of practitioners state that the closure of certain congenital heart defects in children, which had previously required surgery, can now be done through a non-surgical, catheter intervention, with results similar to surgical results. As such, can the non-surgical procedure for defect closure (when properly indicated) be included into the benefit package of BPJS? That clinical question can be answered with regards to the 4 PICO elements:

- **P** *(Patient)* : Child with ABC congenital heart disease;
- **I** *(Intervention)* : Defect closure with catheter intervention;
- **C** *(Comparison)* : Defect closure with surgery;
- **O** *(Outcome)* : Success rate.

Therefore, the problem can be put into the form of research question as follows:

- In children with ABC congenital heart defect, would defect closure with catheterization, compared to surgery, yield a similarly good result?

Other aspects can be questioned separately and specifically. A specific question can be constructed to use keywords in a literature search. For instance, the question of cost comparison of the interventions can be construed as follows:

- In children with ABC congenital heart defect, would catheter closure require less cost compared to surgery?

3. Determining the methods to be used in the HTA process

We should determine the methods for data collection based on the problem. As mentioned, most HTA processes use integrative methods to answer various technical questions, in addition to primary data obtained from Indonesian hospitals. Primary data is a must for economic evaluations.

4. Collecting primary data

Primary data can be obtained from hospital medical records or routine hospital or District Health Office records, etc. In order to obtain primary data, direct interviews with patients or their families are needed to determine treatment and cost. Primary data with good validity can be obtained from formal studies, observational studies, and clinical trials. However, these types of formal studies are rarely used in HTA. The required data for economic analysis must be obtained from local/national data, not only from the literature. In addition, ethical approval is usually needed to obtain primary data.

5. Searching for scientific evidence

Searching for scientific evidence for HTA requires a distinct approach on what is to be learned, trained, or practiced. This topic will be described separately in **Chapter 4.** Unsystematic evidence searching will result in the retrieval of irrelevant "evidence," or, on the contrary, no evidence at all. Information/evidence needs to be collected and synthesized, including:

- The technology itself (drugs, devices, diagnostic kits)
- Safety, efficacy, and effectiveness of drugs, devices, diagnostic and therapeutic procedures
- Indications for use
- Target population for technology use
- Procedure for the use of the technology
- Patient settings (outpatient, inpatient, ICU)
- Service providers (doctors, midwives, nurses)
- Alternative technology as a comparison
- Impact (cost or quality of life)
- Source of available evidence on the use of the technology
- Use of HTA results for clinical guideline, policy, etc.

6. Critical appraisal of the evidence

All scientific evidence used to answer HTA questions should be critically appraised, especially with regards to its **validity**, **importance**, and **applicability**. The details of this procedure will be explained in **Chapter 5**.

7. Synthesis of the critical appraisal results and recommendations

The results of the critical appraisal should be recorded, discussed, and summarized in the conclusions and recommendations. Conclusions and recommendations can be developed for each HTA question, one-by-one, or they can be combined into one distinct chapter. The main point is to communicate with the reader in an effective, informative method.

The best synthesis result can be obtained if the process is conducted according to the standards of systematic review or meta-analysis, such as criteria for the Cochrane Review. **PRISMA (Preferred Reporting Items for Systematic Reviews) is** the suggested guide for meta-analysis for clinical trials, and is the developed version of QUOROM (Quality of Reporting of Meta-analysis). For observational study, MOOSE (Meta-analysis of Observational Studies) is recommended.

8. Conclusions and recommendations

Conclusions can be made with recommendations relevant to the purpose of the study, based on the results of clinical and economic assessments (and others, if any).

9. Disseminating the HTA results

The results of the assessment must be disseminated to other relevant parties. After the final editing, the final draft will be disseminated and commented on by external peer reviewers, i.e., experts on the assessed

topics but not involved in the HTA process. Input and correction are expected to improve the HTA report. The final report should be agreed upon by the HTA Committee before final printing and dissemination.

10. Monitoring the recommendations

The primary task of the HTA Committee is to conduct a systematic and objective assessment on the impact of technology use in the medical field. The final assessment is reported as a recommendation to adopt or not adopt the assessed technology, with or without certain conditions.

The decision on adoption of a recommendation is not the main task of the HTA Committee. However, since the main goal of the assessment is to increase the quality of health service, the HTA Committee (with or without special task) must monitor the implementation of the recommendations. In general, a recommendation to not adopt a certain technology should also be applied, especially if the reasons are concerned with safety of the tech-nology, as well as ethical, moral, or religious issues. However, a recom-mendation to adopt the technology does not necessarily mean that it should be directly implemented, because other factors should be considered, such as costs, facilities, priorities, politics, etc.

Chapter 4
Searching for Evidence

Valid and up-to-date research findings in the literature on the assessed topics are necessary for evidence-based HTA. The internet is the major source of literature evidence. In order to obtain literature findings with appropriate bearing on the topic, skills to access literature from the internet should be improved. Another important criterion is the availability or access to full-text articles. After articles or other evidence are collected, the next step is a critical appraisal of these resources, including the validity of the study, the clinical importance of the results, and the applicability of the results to our patients.

The details on how to search for literature evidence can be obtained from various resources, some of which are explained in great detail, as mentioned in the references. In this short chapter, we will only discuss the principles of internet literature searches. This activity is time-consuming, but with a strategy done in a repeated manner, the scientific evidence can be obtained in a timely way.

Strategy for literature searching

The internet is the best tool to obtain necessary information on the topic at hand. Millions of original articles, systematic reviews, and meta-analyses can be accessed on the internet. Mobile phones are now increasingly sophisticated, so that internet access is available anywhere and anytime.

Many databases are good sources of recent and past data. Hence, databases are recommended as the main source for searching the literature. Here are some examples of commonly-used databases:

- PubMed (www.pubmed.gov)

- The Cochrane Library (www.cochrane.org)

- Clinical Evidence (www.clinicalevidence.com)

- EMBASSE (www.embasse.com)

- CINAHL (www.cinahl.com)

- Highwire Press (www.highwire.org)

- OVID

- Others: Proquest, Science Direct, MD Consult, SAGE, Scopus

The use of keywords

Using keywords is an effective and efficient way to search articles or other scientific evidence. Since most international scientific literature is written in English, the keywords should also use English spelling.

Searching the literature without an appropriate strategy or using un-planned keywords will eventually result in undirected articles or evidence. On the one hand, results may yield enormous numbers of articles that are irrelevant to the topic, or, only a small number of articles, or nothing at all.

Keywords can be obtained easily if the research question or HTA has been developed and consists of the following elements:

- **P**: *Patient / Population / Problem* – i.e. , patients or questionable problems
- **I**: *Intervention /Index / Indicator*, i.e., which intervention to be included in the assessment
- **C**: *Comparison*, i.e., comparison to the intervention
- **O**: *Outcome*, the expected result

Examples:

- In patients older than 60 years, would PSA screening, compared to no screening, detect prostate cancer earlier?

- In patients with prostate cancer, does early diagnosis, compared to late diagnosis, have a better outcome?

- In patients with primary pulmonary hypertension, can oral sildenafil be used to reduce pulmonary pressure?
- Is minimally invasive surgery to cure chronic cholecystitis cheaper than standard surgery?

From those research questions, several keywords (in English) would be selected as terms to be searched on the internet. Note that a disease or medical condition often has more than one or even many synonyms. Therefore, the Boolean operator OR / AND should be used as follows:

- (Population OR synonym1 OR synonym 2……..) AND
- (Intervention OR synonym1 OR synonym 2……) AND
- (Comparison OR synonym1 OR synonym 2…..) AND
- (*Outcome* OR synonym1 OR synonym 2…....)

As many as possible synonyms should be included. The best synonyms come from MeSH (medical subject headings) or a Thesaurus. Alternatively, one can use textwords (look for similarity in the text). In practice, the search should always begin with MeSH keywords. If the word is not found, then the search should be continued with textwords. The "Exp" or "explode" feature is available at Medline and some databases. This feature allows searches in all branches of a certain word or terminology, using a single word. If we search the term "colonic neoplasm ("exp")", then we should include all these terms: **colorectal neoplasm, colonic polyps, colorectal neoplasms, hereditary non-polyposis, and sigmoid neoplasm.**

Boolean operators

Boolean operators (first invented by the English mathematician, **George Boole**, 1815-1864) are simple words that can be used to combine or separate search results on the internet. The most frequent Boolean operators used are **OR, AND,** and **NOT**. The use of the Boolean operator OR includes all articles that contain the term or word. For instance, the words **immuni-zation OR screening** means that all articles with the word **immunization** or **screening** will be included; while the word **immunization AND**

screening means that only articles consisting of the words **immunization** and **screening** will be included; articles which contain the term only **immunization** or only **screening** will be automatically excluded. For the terms **immunization NOT screening**, then articles would have the word **immunization**, but not the word **screening**.

If we perform a comprehensive search with PICO criteria in English, it should be written follows:

- (screen* OR early detection) AND
- (colorectal cancer OR bowel cancer) AND
- (mortality OR death* OR survival)

If we typed all the above wordings, then the search will result in a limited number of articles. Typing only the term "screening test" and "colorectal cancer" would not result in countless outcomes, but in articles of low relevance. The term "screen" will include all words that begin with screen, such as screen, screened, and screening. The symbol (*) has different meanings according to the database used. The general strategy for using Boolean operators is to first increase the sensitivity, then the specificity.

The common formula is as follows:

1. One keyword concerning the patient
2. Another keyword concerning the patient
3. #1 OR #2
4. One keyword concerning the intervention
5. Another keyword concerning the intervention
6. #4 OR #5
7. One keyword concerning the outcome
8. Another keyword concerning the outcome
9. #7 OR #8
10. #3 AND #6 AND 9.

Sensitivity refers to the likelihood of including all the relevant components, while **specificity** refers to the likelihood of excluding all irrelevant com-ponents.

How can the sensitivity be increased? In order to increase the sensitivity level, i.e., obtain more resources or articles, we can do the following:

- Increase the search using a wider terminology from a Thesaurus

- Use the textwords in the database

- Use "truncation and wildcards"

- Use Boolean operators to confirm the inclusion of all alternatives.

A. List of Orders and Terminology in PubMed

ORDER

OR   Finding articles which contain at least one of the words or phrases

AND   Finding articles which contain both words or phrases

NEAR   Include both words, and add 5 more words "near" the aim words. (This order is not available in PubMed.)

NOT   Exclude all articles containing the written words or phrases

Limits The search is limited in several parameters (by date, language, etc)

()   Used to group words

*****   *Truncation.* Act as wildcard that represents all the characters after the term. For example: *vaccine, vaccine, vaccines, vaccination*

[ti] / ti: Title. Finding articles which contain that word in the title.

so/[so] *Source.* Finding articles from certain resources.

MeSH   Medical subject headings. The list of structured words from keywords used in PubMed and Cochrane.

" " Used to give order to the database for searching a specific phrase. If we cannot find the phrase, then the words will be combined with **AND** automatically.

Other databases may use the wildcard "?" and "?*", e.g., gyn?ecology □ gynaecology, gynecology; randomi?* □ randomisation, randomization, randomised, etc.

Limiting the literature search

All databases have features to limit the choices according to certain criteria, such as type of publication, journal, or book, as well as age group, gender, year of publication, setting, and language. In PubMed, the search limit can be used by clicking the term "limit."

Another helpful limitation is to identify the type of article desired: diagnostic, prognostic, therapeutic, preventive, or even harmful etiologies. Other studies may be selected on the basis of the clinical questions. In PubMed, this type of search can be done using clinical queries. Using this method, searches can be performed that prioritize both sensitivity and specificity.

Scientific evidence searches depend entirely on the questions. For interventions, the best evidence would be found in a systematic review of RCTs. The best systematic reviews are available in The Cochrane Database of Systematic Reviews in the Cochrane Library. However, almost all Cochrane Reviews can be found in PubMed. We should always try to find the best evidence available.

Chapter 5
Critical Appraisal of Medical Literature

Introduction

Every scientific article for HTA has to be critically appraised. Critical appraisal should also be done for systematic reviews (SRs) and meta-analyses, even though the individual research reports comprising these articles have already been appraised.

Scientific reports in journals have more or less a uniform format. The article begins with the **Title, Authors** and their **Institutions**, followed by an **Abstract** and **Keywords**. The main text includes the **Introduction, Methods, Results,** and **Discussion**. **Acknowledgments** and the **Conflict of Interest Statement** are also frequently found in articles. The article ends with **References**. The complete critical appraisal should include the whole article, from the **Title** to the **References.**

A short summary of critical appraisal

Comprehensive critical appraisal details can be found in the References. Here is a summary of the main points:

- **Title:** The title should be simple, but represent the main content of the report. The title should not include abbreviations, except for common ones. The title need not be written as a complete sentence with a subject, predicate, and objects. The design of the study is recommended for inclusion in the Title.

- **Authors** and institutions: Use the first name (full or initial), middle name (initial), and family name (full). Every author should include his/her institution's name and affiliation.

- **Abstract**: may be in the form of one paragraph or more common-ly structured abstract consisting of the background, objective, methods, results, and conclusion. The abstract is a brief summary of the most important aspects. Avoid abbreviations. Results should include the clinical findings, not only P values.

- **Introduction:** a simple explanation of why this study was done, the research question or hypothesis, and the aim of the study. The Introduction should be supported by strong references.

- **Methods:** detailed explanation of how the study was done. The design, population, sample, sample size estimation, as well as method of randomization and blinding methods (if applicable) should be described. The analyses, and on which data, as well as programs used should be stated. The approval statement from the Ethics Committee can be included in this section.

- **Results:** begins with a description of subjects, followed by a logical explanation of the answer to the main question, followed by that of the second question, and continuing chronologically. In the Results section, any comments or comparisons to current knowledge on the topic are not necessary.

- **Discussion:** interprets the findings in light of current knowledge on the topic, and connects clinical research with practice. The strengths and limitations of the research should be mentioned. The last paragraph in the Discussion is usually a conclusion of the findings, often with suggestions for follow-up.

- **References:** should be formatted according to the journal's style.

Critical appraisals of HTA may not be done as comprehensively as mentioned above, but should be done according to the guidelines for the main topics. Three parameters to be assessed are:

o   *Validity*

o   *Importance*

o   *Applicability*

or, in short, VIA. Only reports with good validity, clinically important, and applicable results are considered to be valuable input for HTA.

Validity

Validity is mostly assessed in the **Methods** section. Validity includes an assessment of the appropriateness of the design for the objectives, correct subject recruitment, a sufficient number of participants, appropriate allocation, accuracy of the intervention and measurements, as well as correct analysis and interpretation. The **Results** section also contributes to validity, especially for acceptable loss-to-follow up (generally maximum of 20%). Incomplete reports are frequently an obstacle to confirming the validity of a study. For instance, the study may not mention the randomization method, or how participants were selected from the population. Validity should be linked to aspects of treatment, which differ from diagnostics, prevention/therapy, prognosis, and economic factors.

Importance

Importance, in this context, refers to **clinical** importance, in terms of treatment. Importance can be found in the **Results** section. The meaning of importance differs according to type of research design. In a diagnostic study, a 98% sensitivity level for screening trials is important. For clinical trials, the proportional difference in cure rate or the level of certain matter may determine the importance of a certain drug or procedure. In pragmatic clinical trials, the number needed to treat is a gold standard to determine the importance of the result. See page 50. For a prognostic study, the absolute risk (and the relative risk) determine the importance of the results.

We would like to emphasize that **clinical importance** is different from **statistical significance**. If clinically the difference is not important (for

example, if the difference in cure rate between a new drug and the standard drug is only 2%), then the P value is of no use. No matter how small the P value, if there is only a small clinical difference, we would not change our practices. If the results showed a clinically important difference, then we should check the P value to confirm the statistical significance.

Applicability

No matter how valid the methods or how important the results, if it cannot be applied to our patients, then the evidence is useless in our context. For instance, a report showed that a surfactant administered in the neonatal ICU had important effects for the treatment of premature infants with respiratory distress syndrome. But, if a neonatal ICU and surfactants are not available in the clinic or city, then the articles only add to our knowl-edge, but are useless for clinical practice. The first and main question to address applicability is the similarity of our patient with the participants in the study. For example, if our clinical question includes sepsis in neonates, but the assessed articles describe sepsis in youth or geriatrics, then the results cannot be applied to our patients.

Should all articles be evaluated in according to the three aspects (VIA), and in that order? The answer is no. In the context of education or training, the critical appraisal should be done in the order of V-I-A. In practice, however, we may start with applicability. If the facilities to perform a treatment or procedure are not available, then we can discontinue reading the article.

Use of computer software

When we do critical appraisal, a simple calculation is almost always neces-sary. The calculation can be done manually or with the help of calculator, and noted as the **CATmaker** (Critical Appraisal Topic maker). This program can be used for many arithmetical calculations and can be downloaded for free from **www.cebm.net/ catmaker-ebm-calculators/.**

In order to calculate diagnostic and therapeutic values (intention to treat analysis), the two necessary items are software to execute our order by providing a comprehensive result along with its confidence interval, and accurate values to fill in the available cells.

Worksheet for critical appraisal

The worksheet suggested by EBM initiators is helpful in the process of critical appraisal. However, the available worksheet may not be appropriate for all articles. Some authors add more items to be assessed in the worksheet. We attempted to make a simple modification to the common worksheet by adding several aspects to be assessed and eliminating some arithmetical calculations. See **References**.

Chapter 6

Study Designs: an Overview

In health technology assessment (HTA), most of the clinical aspects are assessed using an integrative method. Understanding research design is clearly important because it determines the validity of the study. This chapter briefly describes the common designs used in clinical research.

Types of clinical studies

Research design can be classified into two groups based on the intervention: observational and experimental studies. In **observational studies**, the researcher does not choose which participants receive which intervention; the researcher merely observes, measures variables, classifies, and analyzes data. Observational studies may include case reports, case series, cross-sectional studies (including diagnostic tests), cohort studies (e.g., survival analysis), case-control studies, and meta-analyses.

In **experimental studies,** or so-called interventional studies, the researcher determines which participants receive which intervention (by a randomization process). Experimental studies can be done in laboratories, clinical settings, and even communities. Experimental studies done in a clinic are called clinical trials. The gold standard for a clinical trial is the randomized clinical trial (RCT).

Cross-sectional studies

In cross-sectional studies, variables are measured only once, without follow up. This type of study can be descriptive or analytical, prospective or retrospective. Descriptive cross-sectional studies are also called prevalence

studies. Analytical cross-sectional studies assess the relationship between variables. Independent variables and outcomes are assessed simultaneous-ly, so there is no follow-up. Some examples of descriptive cross-sectional studies include:

- Prevalence of thalassemia in Indonesia
- Clinical and laboratory characteristics of swine influenza patients

Some examples of analytical cross-sectional studies:

- Comparison between total cholesterol level among non-obese versus obese children
- Comparison of underweight neonates among mothers with malaria versus mothers without malaria

The analysis performed depends on the type of data. For numerical data (e.g., comparison of cholesterol level), independent t-test is used for independent data or paired t-test for matched data. For nominal data, Chi-square test can be applied to independent data, while McNemar test can be used for matched data.

If the variables are risk factors (existing or not) and the outcome has a two-nominal value (sick or not), then the prevalence ratio can be calculated using the same calculation as relative risk in cohort studies.

Cross-sectional studies are often used to evaluate several risk factors of certain outcomes, therefore, multivariate analysis can be applied. If all risk factors and outcomes are numerical variables, then multiple regression would be applied. For nominal variables, logistic regression would be used instead. Diagnostic studies are a special cross-sectional study frequently used in HTA.

Case-control study

A case-control study aims to assess the role of risk factors in a certain disease. The study starts with recruitment of a group of participants with a

certain disease (case), and a group of participants without the disease (control). See **Figure I-1**. In both groups, suspected risk factors are assessed retrospectively. The frequency of exposure to risk factors in the case group is compared to that of the control group. Examples:

- A study evaluating the relationship between maternal consump-tion of herbal drinks in the first trimester and the incidence of a certain heart disease in infants

- The effect of extra-high voltage (SUTET) on the incidence of malignancy

Case-control studies can be used to investigate rare cases. The correlation between risk factors and effect variables is shown by the odds ratio.

$$Odds = probability / (1\text{-}probability)$$

Therefore, if probability for outcome equals to ¼ (0.25), then the odds for outcome is 0.25.

$$Odds = 0.25/ (1 – 0.25) = 0.25 / 0.75 = 0.33$$

The odds ratio is a comparison of two odds.
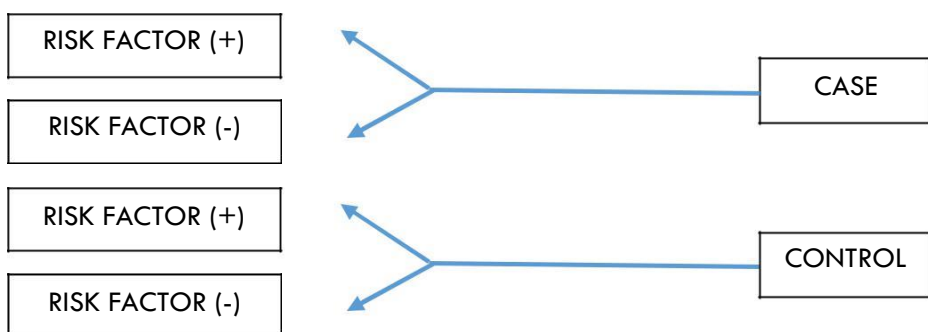
Figure I-1. Scheme of a simple case-control study.

Let's take for example, a case-control study to identify the role of risk factor X on the incidence of chronic renal failure (CRF). See **Figure I-2**. The researcher recruited 50 CRF patients (case group) from one large hospital, and 50 participants without CRF (control group) from the same population. The research would assess the number of participants in both groups who have risk factor X.

The data show that out of 50 cases, 20 had factor X (odds for CRF in the case group = 20/30); while in the control group, only 5 subject had factor X (odds for CRF in the control group = 5/45). The odds ratio (OR) is the comparison between the two odds; in this example OR = 20/30 : 5/45 = 6 (95% confidence interval: 2.3 to 12.7). This means that the participants with factor X have 6 times higher risk of having CRF than participants without factor X. In the population, 95% represented in the samples had risk between 2.3 and 12.7 times.

	Case	Control
Risk (+)	20	5
Risk (-)	30	45
	50	50

Figure I-2. Standard analysis in a case-control study.

This 2 x 2 table shows the result of a case control study. The odds of having CRF in participants with factor X = 20/30, while the odds of having CRF in participants without factor X = 5/45. The *odds* ratio (OR) = 20/30 ÷ 4/45 = 6 (95% CI: 2.3 to 12.7). (CRF = chronic renal failure, CI = confidence interval)

Cohort study

A cohort study is a research design to test a hypothesis about potential risk factors or etiologies of a disease or medical condition. Selection of individ-uals for cohort studies is done from the same population. Selection can be done in one of two ways:

- **Cohort design with internal comparison:** a group of participants without risk factors is recruited and followed. Some will be exposed to the potential risk factors, and the others will not.

- **Cohort design with external comparison** *(double cohort study)*: groups of participants with and without exposure to risk factors are recruited from the same population. These two groups (ex-posed and unexposed) will be followed until the effect or disease appears. A cohort study is the only design which shows the inci-dence of a certain disease or defect.

A cohort study may be prospective or retrospective. Examples of pros-pective cohort studies:

- Effect of noise on the incidence of deafness among laborers

- Incidence of post-stroke transient ischemic attack (TIA)

- Incidence of malignancy in infants conceived by IVF.

In a cohort study, descriptive data, e.g., incidence of a disease or defect, can also be expressed as relative risk, i.e., the number of participants with risk factor exposure who develop the effect compared to the number of partic-ipants not exposed to the risk factor but who develop the effect. See **Figure I-3**. Retrospective cohort studies have basic principles similar to pro-spective cohort studies. Predictors/risk variables are measured in a group of monitored participants, but the period of follow-up and the measure-ments were done in the past. In order for a retrospective study to have good validity, all past records should be complete.

	Disease +	Disease -	
Risk (+)	a	b	a+b
Risk (-)	c	d	c+d

Figure I-3. Basic analysis of a cohort study.

The risk of developing disease in the group with risk factors = a/(a+b). The risk of developing disease in the group without risk factors = c/(c+d). Relative risk = a/(a+b) ÷ c/(c+d).

Diagnostic test

The majority of diagnostic tests have a cross-sectional design. In diagnostic tests, no intervention will be conducted or measured. A group of partici-pants is assessed using two types of examinations. The first test is the one to be studied, while the second one is the gold standard test for that disease. Then results are analyzed.

The characteristics of diagnostic tests are similar to prognostic tests. The difference is merely the outcome variables. For instance, a diagnostic test aims to identify if severe head trauma is predictive of the incidence of intracranial bleeding, while a prognostic test aims to understand if severe head trauma is predictive of mortality due to intracranial bleeding. In this example, there is a clear difference between designs. The difference lies in the outcome variables. The diagnostic test predicts the existence of the disease, while the prognostic test predicts the outcome of the disease.

A diagnostic test does not always correctly identify (positive) all partici-pants with the symptoms. Likewise, diagnostic tests may not have a nega-tive result in all participants without the symptoms. An ideal diagnostic test is rare. Indeed, almost all diagnostic tests have a probability of a positive

result in participants without the disease (false positive), and vice versa, a probability of a negative result in participants with the disease (false negative). Furthermore, a good diagnostic test should be safe, simple, not harmful/invasive, reliable, affordable, and yield quick results.

The structure of a diagnostic test is similar to observational studies, in that it has a predictor variable (the test result) and an effect/outcome variable (with/without disease). The predictor variable can be classified into a dichotomous scale (positive or negative), categorical scale (+++, ++, +, -), or numeric scale (milligrams per deciliter).

If the diagnostic test result is categorical or numeric, then a cut-off point is necessary to differentiate participants into with or without disease. The outcome variable in a diagnostic test is the presence or absence of the disease, based on a gold standard.

Even though the structure of the diagnostic test is similar to that of an observational study, the analysis is very different. An observational study has the aim of assessing an etiology or risk factors, while a diagnostic test has the aim of differentiating participants with the disease from those without disease. The result of diagnostic test can be summarized into a 2x2 table, consisting of cells a, b, c, and d.

- **Cell a** contains the number of subjects with the disease, based on the gold standard and a positive test result (true positive, TP).

- **Cell b** contains the number of participants without the disease based on the gold standard, yet diagnosed as having the disease based on the test (false positive, FP).

- **Cell c** has the number of participants with the disease based on the gold standard, yet diagnosed as not having the disease by the test (false negative, FN).

- **Cell d** contains the number of participants without the disease based on the gold standard and also diagnosed as not having the disease by the test (true negative, TN).

Generally, the diagnostic test analysis includes (see **Figure I-4**):

- **Sensitivity**: the proportion of sick participants with positive test results (TP). The sensitivity shows the capability of the diagnostic test to detect the disease. Sensitivity = TP/(TP+FP)

- **Specificity**: the proportion of subjects without the disease with negative test (TN); it shows the capability of the diagnostic test to identify subjects without disease. Specificity = TN/(TN+FN)

- **Prevalence,** or so-called prior probability: probability of an individual (based on demography and clinical characteristics) to develop a disease before running the diagnostic test. Prevalence = (TP+FP)/(TP+FP+FN+TN)

- **Positive predictive value**: the probability of an individual to have a positive diagnostic test result and the disease, based on the gold standard. The positive predictive value = TP/(TP+FP).

- **Negative predictive value**: the probability of an individual to have a negative diagnostic result and not have the disease. The negative predictive value = TN/(FN+TN).

- **Likelihood ratio**: indicates the possibility of participants with the disease having a certain result, divided by the possibility of participants without disease who have a similar result. **Positive likelihood ratio** is a comparison between the proportion of participants with disease and a positive result to the proportion of participants without disease, but with a positive result. The formula is as follows: **sensitivity / (1-specificity)**. Conversely, the **negative likelihood ratio** is a comparison between the propor-tion of sick participants and a negative result to the proportion of participants without disease and with a negative result. The formula is as follows: (**1-sensitivity) / specificity.**

Note: In the interpretation of diagnostic test, this acronym can be used:

- *SnNOut = A very sensitive test, when negative, rules out diagnosis.*
- *SpPIn = A very specific test, when positive, rules in diagnosis.*

Many of the above values can be obtained in the diagnostic test, but in practice, the positive and negative predictive values are most often used. For instance, if we get a positive diagnostic test result, we must ask, "What is the chance (percentage) that the participant has the disease?" The answer lies in the positive predictive value. On the contrary, if the result was negative, then the negative predictive value would tell us the probability that a patient was truly healthy (without disease). Both positive and negative predictive values are heavily influenced by the disease prevalence. For that reason, before using a new diagnostic test reported in a journal, we must note if the disease prevalence in the reported study is similar to that in our population.

Gold standard

	Disease +	Disease -	
Test (+)	a	b	a+b
Test (-)	c	d	c+d
	a+c	b+d	a+b+c+d

Figure I-4. The table indicates the diagnostic test result. Based on this table, the following values can be calculated:

Prevalence	= (a+c)/(a+b+c+d)
Sensitivity	= a/(a+c)
Specificity	= d/(b+d)
Positive predictive value	= a/(a+b)
Negative predictive value	= d/(c+d)
Likelihood ratio of positive test	= a/(a+c) : b/(b+d)
Likelihood ratio of negative test	= d/(b+d) : c/(a+c)

43

Experimental study

An experimental or interventional study can be considered as a cohort study in which the researcher "manipulates" the predictor value or risk factors through certain interventions and then analyzes the effect or outcome variables as the result of the intervention. In contrast to an observational study, an experimental study is able to show a causal relationship between the predictor and the effect variables, and is the best design to control for effects of confounding variables.

Experimental research can be classified into 3 groups:

- A true experiment with main criteria, i.e., randomization

- A quasi-experiment, without randomization

- A pre-experimental design, generally without a control group

Clinical trial

A clinical trial is an experimental study to evaluate the effect of a drug or medical procedure in humans. Clinical trials for drugs have four phases: phase 1 to mainly assess safety; phase 2 to assess the initial pharmacological effect; phase 3 to evaluate all aspects comprehensively; and phase 4 to provide surveillance after the drug has been used in practice. It is implied that a clinical trial has a prospective research design, with an intervention and human participants. The gold standard for clinical trials in this phase is the **randomized controlled trial** (RCT). This type of research design has the highest priority, in terms of level of evidence.

A clinical trial study design can determine the effectiveness of a certain drug or therapeutic procedure. Most commonly, the effect of a certain drug is compared to the effect in a control group (without a drug, or with a placebo or the standard drug of choice). The result is based on the difference in outcomes between the intervention and control groups.

Three comparability criteria in a clinical trial

In daily practice, we would treat a patient with disease A by giving drug B, and ask him/her to follow up one week later. If all the symptoms and signs are missing, then the patient would be considered recovered. The question is: was the recovery merely due to drug B? The answer is no, not necessarily. Besides the drug intervention, three other conditions can cause patients to recover or otherwise be healed.

- First, recovery may occur because of the natural pathway of the disease. With or without the drug, the patient would reco-ver in one week (natural history of the disease or prognostic factors);

- Second, the patient may have consumed other drugs or herbal formulas, made a change to his diet, or had adequate rest, etc. (extraneous factors);

- Third, the criteria for recovery or outcome used was imprecise or subjectively assessed (outcome measurement factor).

In order to make a fair comparison of intervention results between the experimental group (E) and the control group (C), the three conditions mentioned above should be similar or comparable between groups.

- **Comparable prognostic factors**. The E and C groups should be similar in terms of prognostic factors. For example, one group cannot have more patients with severe disease, higher mean cholesterol level, older age, worse nutritional status, etc., compared to the other group. In order to obtain 2 compar-able groups, the patients need to be randomized. Randomi-zation tends to divide patients with prognostic factors and the confounding variables equally between the two groups.

- **Comparable intervention**. Participants in both groups should be treated equally, except for the use of the drug or procedure

under investigation. As such, the participants in E group should not receive better attention, better health facilities, or receive additional diet supplements or drugs, while partici-pants in C group do not. In order to guarantee that the groups receive similar treatment, **masking/blinding** should be done. In blinding, one or more of the relevant parties in the clinical trial (researchers, participants, evaluators, lab assistants, etc.) do not know the type of therapy given. If double-blinding is possible (both researcher and participants do not know the treatment), then the validity of the study would be excellent.

- **Comparable outcome measurement.** If the clinical outcome is "hard data" such as died or survived, or the laboratory test is done on a standardized automatic machine, then in the context of outcome measurements the blinding is not (highly) necessary. However, if the outcome is subjective (pain, anxiety, and the like) or requires interpretation of an exam (ultrasound or X-ray), then blinding is extremely important.

If in a clinical trial, the three comparable criteria are equally applied (through **randomization and blinding**), then the difference between the outcomes in the E and C groups should be due to the intervention. There-fore, the best design for a clinical trial is the randomized, double-blind, clinical trial. If the number of participants is sufficient, then randomization can be done. However, blinding may not always be possible, for instance, in a clinical trial comparing the effectiveness of a certain drug to surgery for a certain disease or medical condition.

The validity of a clinical trial is also determined by the participants com-pleting the study until the end (**completion of follow-up**). In general if the number of participants who completed the study is less than 80%, then the validity of the clinical trial is questionable.

Most clinical trial outcomes are numerical (e.g., cholesterol level, body weight, blood pressure, etc.) or dichotomous nominal variables (e.g., died or survived, recovered or not).

Pragmatic and explanatory clinical trials

Clinical trials are done in human patients. There are two types of clinical trials: the pragmatic trial and the explanatory trial. The pragmatic trial yields relevant results which are directly applicable to practice. Some characteristics of a pragmatic trial include:

- A similar spectrum of patients who meet the not-so-strict inclusion criteria. For instance, in a clinical trial for diabetic drugs, the inclusion criteria may be general diabetic patients. Some of these patients might be obese, malnourished, hypertensive, hypercholesterolemic, etc. If the inclusion criteria are too tight (e.g., diabetic patients without obesity, without hypertension, without hyperlipidemia, no history of coronary heart disease, etc.), then while the internal validity would be extremely good, the external validity would be low, as the applicability of the result would be limited in practice.

- A clinical outcome is the goal of a pragmatic trial, without con-sidering the mechanism of how that outcome occurred. For example, if a clinical trial is done to assess if a traditional drug can stimulate appetite in children, the most important result is the actual outcome (e.g., increased appetite that was objectively measured as increased body weight), not the mechanism of how the appetite was increased.

- If the result is binomial (recovered or not, success or failure, etc.), then the study is called an intention-to-treat analysis. The salient feature is that all the randomized participants are included in the final analysis, regardless of completion of the study. (See below).

The other type of clinical trial is called an explanatory trial. Its aim is to explain the mechanism or why there are outcome differences between two types of medication. This type of clinical trial is extremely useful for scientific understanding, but not as relevant in daily practice.

In an explanatory trial, the final analysis only involves participants who follow the research until the end (per protocol analysis or on treatment analysis).

Analysis in a pragmatic trial

As mentioned previously, a valid clinical trial should have a random-ization process. Let's design a hypothetical trial. The effectiveness of experimental drug (E) is to be compared to the standard drug (C). The sample size calculation result is a minimum of 80 patients per group. The researcher designs a pragmatic clinical trial with 160 total participants. The patients are consecutively recruited, and every subject is randomized to receive either the experimental drug (E group), or the standard drug (C group). After the study is completed, out of 80 participants in the E group, 60 recovered, 15 failed, and 5 were lost to follow up. In the C group, 45 participants recovered, 25 failed, 2 moved to the E group, and 8 were lost to follow up.

The conventional method is to analyze the data by making a 2x2 table, involving only participants who completed the study, then calculating the p value. The weakness of this method is that it ignores the partici-pants who were unable to complete the research or were lost from the observation, even though the randomization procedure created two equal groups, 80 participants in the E group and 80 participants in the C group. If number of lost to follow up patients is ignored, i.e., not included in the analysis, then the two groups would no longer be equal.

A second weakness can occur in making an analogy in practice. If we treat 10 patients with a certain disease, and one week later 8 patients are recovered, then we cannot say that our treated patients recovered 100%. Therefore, a new analysis method with heavy emphasize on clinical aspects can be introduced, the intention-to-treat analysis. Using this method, all included participants, usually along with the randomized ones, are included in the analysis. Participants who were lost to follow up are considered as failure, but will be kept in the initial group. Com-

pare the following two 2x2 tables (**Figure I-5** and **Figure I-6**): the first is the per protocol analysis, the second is the intention-to-treat analysis.

Per protocol analysis

The number of initial participants was 80 each in the E and C groups. In the E group of 80 participants, 60 recovered, 15 failed, and 5 were lost to follow up (LTF). In the C group of 80 participants, 45 recovered, 25 failed, and 10 were LTF. In this analysis, the LTF participants are ignored. **Figure I-5** shows that the result show statistically significant.

	Recovered	Failed	Total
E	60	15	75
C	45	25	70
Total	105	40	145

Figure I-5. Per protocol analysis of a clinical trial.

Statistical analysis yielded a p value of 0.047 (statistically significant).

Intention-to-treat analysis

In the case of intention-to-treat analysis, all randomized participants are included in the analysis. In the 2x2 table, participants lost to follow up are considered to be not recovered and are retained in their initial randomized group. Therefore, the E group has 60 recovered and 20 failed patients (total of 80 patients). In the C group, there are 45 recovered patients and 35 failed patients (total of 80 patients). Based on the table, the following values can be calculated: control event rate

(CER), experimental event rate (EER), relative risk reduction (RRR), absolute risk reduction (ARR), and the number needed to treat (NNT). See **Figure I-6.**

	Recovered	Failed	LTF	Total
E	60	15	5	80
C	45	25	10	80
Total	105	40	15	160

Figure I-6. Intention-to-treat analysis. LTF = loss to follow-up

- Control event rate (CER) = 35/80 = 44% = 0.44
- Experimental event rate (EER) = 20/80 = 25% = 0.25
- Relative risk reduction = (CER-EER)/CER=(0.44-0.25)/0.44= 0.43
- Absolute rusk reduction = CER - EER = 0.19
- Number needed to treat (NNT) = 1/ARR = 1/0.19 = 5

Notes

In the above calculations, based on convention, an event is defined as failure. Therefore, the control event rate (CER) means the proportion of failure in the control group; while the experimental event rate (EER) is the proportion of failure in the experimental group.

Meta-analysis

The aim of a meta-analysis is to integrate the results of several individual studies into a one large analysis, in order to increase our understanding of the item of study. A meta-analysis includes retrospective observation-al studies, and the participants are articles or original research reports, either published or unpublished.

Meta-analysis is a method to quantitatively combine relevant studies through a systematic review, or in short, meta-analysis is a systematic review that is analyzed statistically to obtain one combined result. To date, the most common meta-analysis design used is the clinical trial.

There are 5 steps to developing a meta-analysis,:

1   Formulating the research problem, which is essential in any type of study design

2   Identifying relevant studies, published or unpublished

3   Determining the inclusion and exclusion criteria for individual studies, including the study design, participants' charac-teristics, minimum sample size, dosage, setting, language, etc.

4   Abstraction and data weighing from each study, as a study with a large number of participants and good methodology should receive greater weight.

5   Data analysis aims to combine various study results to obtain one pooled result. If the result is numeric, then the pooled result would be a standardized mean difference, i.e., mean difference divided by the standard deviation. But if the result is binomial, then the pooled result would be odds, incidence, risk difference, odds ratio, or relative risk.

Since the sample size of each individual study is different, in order to combine many individual studies, a certain statistical technique is used. The most common techniques are the fixed effect model and the random effect model. In the fixed effect model, the variability between studies is ignored while intra-study variability is based on the factor of chance. Using this technique, we can obtain a narrow confidence interval. In the random effect model, both intra-study and inter-study variability is considered, therefore, the confidence interval would be wider.

In general, if the combined studies are homogenous, then we would use the fixed effect model; but, if they are heterogeneous, then the random

effect model would apply. Often times, both models are applied to show that the results are not much different.

As in other study reports, a meta-analysis starts with an introduction of the necessity of the meta-analysis, followed by an explanation of how the study will be conducted, the search strategy, key words, database, inclusion and exclusion criteria, language, year of publication, etc.

The result will be reported in narrative form and almost always accompanied with figures called a forest plot. A complete forest plot consists of the following components. See **Figure I-7**:

- **Title or label** to show the type of investigated drug (predictor variable) and the expected outcome variable

- **Identity** of each study in the meta-analysis

- **Data from each study**, consisting of the total number of participants and number of participants with the outcome, both in the experimental (E) and control (C) groups

- **The final result of the statistical calculation** in each study, such as the standardized mean difference for numerical variable, or odds ratio, risk difference, or relative risk. For nominal variables, the confidence interval (CI) is given. In addition to numerical values, the results should be illustrated in a square figure with a horizontal line that describes the CI.

- **Weight percentage** of each study, stated as a number or presented as the area length of the square. If the weight is larger, then the area length will be bigger.

- The **horizontal line** at the base of the forest plot shows the measurement scale of the treatment effect. This scale should be read thoroughly, as the left side does not always indicate better treatment effect.

- The **vertical line** in the middle of the forest plot represents a similar effect result between the intervention and control groups, or the so-called **no-effect line**.

- **Pooled results** are described in a diamond shape. The middle of the figure indicates the point estimate, and the farthest distance between the left and right ends shows the confidence interval. If this confidence interval line goes through the no-effect line, then we can conclude that there is no difference in the outcome variables between the intervention and control groups. The pooled result can also be represented in numbers in addition to figure.

- The assessment of **heterogeneity** is defined as the Chi-square test result, p value, and I_2. A p value more than 0.05 indicates that there is no significant heterogeneity, while I_2 less than 50% also shows that studies included in the meta-analysis are not heterogeneous.

Study	E: n/N	C: n/N	OR (95% CI)	Weigh (%)	OR (95%CI)
1	5/21	5/22		8.9	1.06 (0.54 to 1.52)
2				25.2	0.41 (0.25 to 0.98)
3	13/109	27/111		45.2	0.42 (0.34 to 0.94)
4	14/50	19/51		20.7	0.65 (0.48 to 0.87)
Total (95% CI)	242	244		100.0	0.52 (0.46 to 0.92)
Heterogeneity test: x_2=2.62, df=2, I_2=0.0%			0.01 0.1 0 10 100 *E better* *C better*		
Overall test: z=2.31; P=0.021					
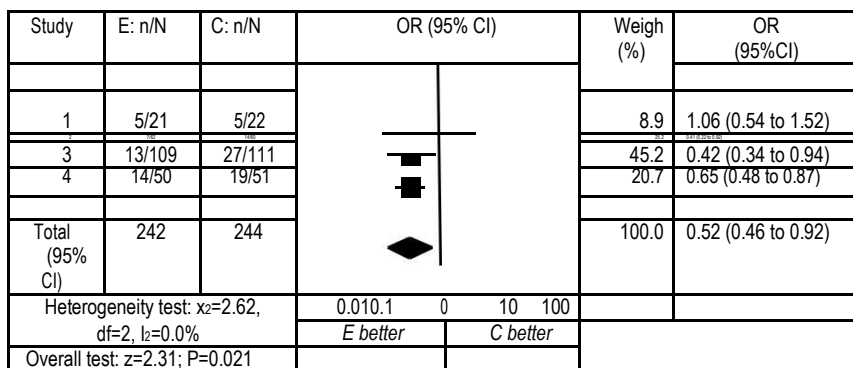
Figure I-7. Forest plot diagram showing a comprehensive result of a meta-analysis.

The included studies specify the number of participants for the E (experimental) and control (C) groups. The number who develop the effect in each group is determined, therefore, an odds ratio can be calculated using its confidence interval. The weight and heterogeneity analysis (p value, I_2) are also included.

53

Publication bias in a meta-analysis

One should take note that publication bias can occur in a meta-analysis. Researchers tend to submit more research for publication, if the results are positive (significant p value). In addition, editors are usually more willing to accept articles with significant results. Studies with insignificant results are often reported in local journals, while the significant ones are reported in international journals. Therefore, publication bias might happen because the literature is dominated with statistically significant study results.

Chapter 7
Social, Ethico-legal, and Religious Aspects

The use of health technology has a broad impact on society, with regards to ethical, legal, social, cultural, and religious values. Some view technology as neutral, free of value. How an individual applies that technology gives it value. Therefore, one might assume that if the aim of the technology is to improve health, then no social, ethical, or legal problems would arise, as long as the technology application has a positive impact on patient health.

In fact, reality is not that simple. These non-technical aspects must be studied, especially with regards to integrating them into HTA. In general, most HTAs have not taken the legal, moral, ethical, or cultural aspects into account. Moreover, these aspects vary widely between countries, so it is difficult to make generalizations. Countries, or even areas within a single country, may have different views on a certain technology application, especially if one considers a sophisticated technology such as stem cell therapy or nanotechnology.

The social and ethical aspects of the health technology application have similar scope of impact on the patient's overall treatment. For example, life support systems (ventilator to assist breathing, pacemaker to support the circulation, etc.) can pose an ethical issue with regards to when to start, when to end, with or without family permission, etc. Also, vaccines or certain drugs may be rejected if they have ethical implications. Although social and ethical aspects are inseparable from HTA, the assessment method of those aspects is still underdeveloped. Some parties have even questioned the definition of assessing the social and ethical impacts of certain health technologies.

In most HTA, social, legal, and psychological aspects may not always involve experts in these areas. However, for technologies that may have broad effects in those aspects, expert opinions are necessary. Also, laypersons may be invited in certain HTA cases. They can give input to the team about their perceptions of a certain health technology.

The factual evidence received from patients about the use of a technology is best obtained from a qualitative study of patients as the user. A good qualitative study can capture patient perceptions and opinions on a certain technology, without the need for statistical analysis, which may not always align with patient views in a qualitative and normative fashion.

In a qualitative study, patients and providers can express their opinion on the use of certain technology. Similarly, the religious implications of certain technologies, which are barely discussed in Western literature, and cultural aspects should be included in the HTA. We, in Indonesia, need to pay special attention to the religious aspect, as the majority of Indonesians are muslims who follow specific regulations in many aspects of life. In the Hadist, the Prophet Muhammad saw. said that Allah will not send any disease unless He provides the medication, therefore, use the medication. However, the Prophet Muhammad saw. prohibited illicit goods as medica-tions, for example, alcohol or items containing pork.

With the advance of technology, religious views, including Islam, must be taken into account in the HTA. Imagine the following recent technologies of breast milk banks, sperm banks, the use of cadavers for transplantation, stem cell technology, nanotechnology, sex reassignment surgery, or pregnancy termination if the fetus is confirmed to have certain disease or syndromes that is incompatible with life. As such, religious leaders should be involved in HTAs that have religious implications.

Chapter 8
Hospital-based HTA

The HTA nationwide study results will be partially applied in hospitals, in the form of clinical practice guidelines corresponding to each hospital, for the treatment of inpatients as well as outpatients. The relationship between research, health technology, clinical practice guidelines, and quality control has been discussed in **Chapter 1.**

In addition to applying nationwide HTA results, hospitals (especially large ones) may require an internal HTA to answer problems specific to that hospital, if no nationwide HTA exists. For example, a large academic hospi-tal would like to assess treatments for acute limb ischemia, to determine if non-surgical intervention is as good as, worse, or better than surgery, in terms of effectiveness and cost.

The condition is specific to a large hospital with comprehensive subspecial-ists. Another example would be an HTA that compares a minimally invasive surgical procedure to the conventional procedure for certain cases. Lower ranked hospitals can also do their own HTA if a national HTA does not exist, for topics specific to that hospital.

In this chapter, we will briefly describe the benefits of applying HTA results in hospitals and an overview of HTA in hospitals.

Benefits of HTA in increasing the service quality in hospitals

With the above background, we can understand that HTA has a large role in hospital service, especially for providing input to the clinical practice guideline (CPG).

Constantly rising quality standard

Many theories have be used to evaluate the quality of health service, from the simplest to the most complex. A health facility or hospital delivers health service to the public, with the final outcome being the **service value**. The health service value is not the same as the service quality. The value is connected to cost, in a broad sense. Mathematically, the value can be described as quality or cost. On the other hand, service quality is represented by the service outcome, i.e., decreased morbidity and mortality, increased quality of life, and patient satisfaction. Good service quality increases the health of the population. In a broad sense, the service quality is considered good if it requires a high cost (e.g., in the United States), but its service value is below that requiring a lower cost (e.g., in the North European countries).

The health service quality should increase over time. Many factors affect service quality, but the most obvious are the fast development of science and technology and the demands from society. Efforts to improve service quality are done, but the majority are carried out in individual sectors, not in a well-planned comprehensive effort. Hospital directors, professionals, nurses, and support systems tend to work individually, without firm coordination. In response to these conditions, the "**clinical governance**" concept was developed to improve service quality. Clinical governance is defined as "a framework through which NHS organizations are accountable for continuously improving the quality of their services and safeguarding high standards of care, by creating an environment in which excellence in clinical care will flourish." Clinical governance is generally comprised 6 elements:

- Education and training,
- Clinical audits,
- Clinical effectiveness,
- Research and development,
- Accountability, and
- Risk management.

These six elements are interconnected to each other and may even overlap. The use of health technology plays an important role, mainly in the clinical effectiveness element, since HTA provides significant input to the development of the clinical practice guideline. A consequence of the use of clinical practice guidelines is quality control. Quality control may lead to questions on whether the clinical practice guideline was followed properly, which can be ascertained through a clinical audit. The implementation of clinical guidelines has a risk of developing unwanted effects. As such, we briefly describe the clinical practice guideline, clinical audit, and risk management.

Clinical practice guidelines

In the era of evidence-based medicine, a clinical practice guideline should be developed based on scientific evidence. In developed or developing countries with limited geographic conditions, national clinical practice guidelines can be used in hospitals nationwide. If more complex services are required, a referral system, including facilities and required resources, is available. However, conditions such as these do not apply to Indonesia, because the country consists of tens of thousands of islands, more than 2,300 of which are inhabited. There is a significant gap in terms of facilities and resources between hospitals/health facilities in large cities and remote areas, especially in the eastern part of Indonesia.

Therefore, two guidelines are needed to bridge the gap in Indonesia: a national guideline, the National Guideline of Medical Service (NGMS), and a clinical practice guideline (CPG), which is considered to be a NGMS but adapted for local conditions, or made with reference to other resources. Note the following explanations:

1   **National Guideline of Medical Service (NGMS) –** is a statement, developed systematically, based on evidence, in order to help doctors (and other service providers) handle specific clinical problems. The NGMS was developed by experts organized by the Ministry of Health, and contains specific recommendations

and "ideal" scenarios, based on the recent developments. The NGMS is necessary for diseases or clinical problems with high volume, high risk, high cost, and high variability.

2   That NGMS needs to be adapted to local conditions for the service facilities in the CPG (**Clinical Practice Guideline**).

In certain hospitals, the suggested HTA results may not be feasible. As such, hospitals, especially lower ranked hospitals, should have CPGs modified to their capabilities.

In HTA, other instruments might be necessary to clarify or describe in detail:

- **Algorithm** – usually required for acute treatment, such as in the emergency ward or intensive care unit (ICU)

- **Clinical Pathway** – necessary for diseases or clinical conditions which require a multidisciplinary approach and have a predictable natural history of disease (e.g., non-hemorrhagic stroke, dengue fever without shock). The clinical pathway of certain interventions, such as Caesarean surgery, appendectomy, closure of congenital heart defect with a device, can also be made since these interventions require multidiscipline approach and have many predictable factors in the disease natural history.

- **Protocol** – a process (complex one) to execute a certain clinical service, e.g., the protocol to install a ventilator or the protocol for hemodialysis.

- **Procedure** – a step-by-step, technical method to do a certain technical task, e.g., lumbar puncture or installing an umbilical catheter.

- **Standing orders** – a consistent order to the nurse to perform a certain intervention while the doctor is away, e.g., give paracetamol to a child with high fever or rectal diazepam to a baby with febrile convulsions

Note that NGMS is only made for a certain number of conditions which fulfill the following criteria (one or more): high volume, high risk, high cost, and high variability in practice.

Therefore, the CPG may vary, but most should be developed in reference to other resources (recent literature, systematic review or meta-analysis, clinical practice guidelines from other countries, guidelines from professional organizations, or MoH program guidelines.)

With the above taxonomy, the term **standard operating procedure** (SOP), is no longer used in the level of service for specific processes or procedures.

HTA should be prepared by professionals in the department or division under the guidance of the Medical Committee, and take effect after being signed by the Director or the heads of the service facility. Every HTA should contain a disclaimer that it is only a recommendation, and may not be applicable to all patients with the condition. This disclaimer should be emphasized so there will be no misunderstanding because:

- CPG is developed for general patients;
- CPG is developed for a single disease;
- Patient response to diagnostic or therapeutic procedures may vary;
- ÇPG is valid when it is printed out; and
- Modern medical practice requires the involvement of the patient and family in clinical decision-making.

In cases when the doctor veers from the CPG recommendations, then he/she should explain the reason in the medical record. If not, the doctor may be considered guilty of malpractice.

Clinical audits

The term **audit** includes a number of activities ranging from unstructured self assessment to a comprehensive review of structures, processes, results,

or impacts. Previously, these activities were called a medical audit, but, in fact, they should be called a clinical audit, because the word 'medical' refers only to doctors, while the service involves other professions.

The process begins with selecting the topic, which may be an indication that the outcome of an event was worse than expected. For instance, the mor-tality of patients with disease X should be nearly zero, but in the past year mortality increased. Once the topic is selected, the medical records of patients with disease X are collected. The audit team discusses the criteria (the variable to be measured) and the standard (the percent to be fulfilled). For example, patients should be examined by DPJP before 9 AM (the criteria), with a standard of 90%. The audit team then verifies the actual percentage that was done. Each aspect that has been established to have criteria and standards, is audited as to the percentage that was fulfilled. This process will reveal the level of HTA. The clinical audit should conclude with recommendations for improvement, including methods, performed by whom, and when to complete them. After some time, there should be a re-audit to assess the implementation of the recommendations.

A clinical audit is considered to be the heart of clinical governance, since through this process, we can assess if the clinical practice guideline (CPG) was conducted properly and if the CPG is sufficient, or requires revision. Hence, the aim of the clinical audit is to improve service quality.

A clinical audit has to be done in a transparent way by a multidisciplinary team, involving all the relevant parties. An audit cannot be transferred to other people such as a professional audit team. In addition, external parties cannot be used as consultants. As such, the auditing process should be conducted by the team that runs the service.

A clinical audit is not considered to be research with a complicated statis-tical analysis. Nor is it a procedure to determine wrongdoing or a discipline mechanism. With this understanding, an audit can be done continuously and involve all aspects, in order to improve service quality.

Patient safety

Risk management, as part of clinical governance, is currently considered to be similar to patient safety. The use of health technology, while expected to have a positive impact, also has the potential to stimulate or create un-wanted effects.

Patient safety is the main topic in the report "**To err is human**" by the Institute of Medicine. Two reports from Colorado & Utah Hospital as well as New York Hospital mentioned that adverse events (AE) occurred in 2.9% and 3.7% of inpatients, respectively, and mortality occurred in 6.6% and 13.6%, respectively, of the total patients with AEs. More than half of the deaths could have been prevented. If those numbers are extrapolated to 33 million inpatients over the United States per year, then every year there are 44,000−98,000 patients die due to medical errors annually. Astonishingly, these numbers are greater than the number of deaths due to traffic acci-dents, or HIV, or breast cancer.

Medical errors that should have been prevented include the following:

- Medication was given to the wrong patient because staff did not pay attention to the patient's identity
- Surgery on the wrong organ or the wrong side of the body due to an error in patient identification
- Wrong label on a blood specimen or other matter
- Wrong diagnosis
- Low infection control in the hospital
- Inaccurate record of drug dosage
- Unclear instruction or using non-standardized acronyms
- Improper function of various devices
- Unauthorized actions
- Staff fatigue

Many events can potentially have negative impacts (unwanted or near-miss events) due to simple things: carelessness, illegibly written instructions, using non-standardized acronyms, not washing hands properly, or storing drugs with similar names in the same location.

The problem of patient safety has gained worldwide attention by the WHO and ministries of health all over the world, including the Indonesian MoH, which has established patient safety measures.

Patient safety is also part of hospital accreditation. Some preventive actions, such as *failure mode and effects analysis* (FMEA), availability of a *rapid response team* (RRT), and searching when the event happens *(root cause analysis)*, *fishbone analysis,* etc., have been put into place.

The key to patient safety in practice is the willingness of all parties to report any unwanted, unexpected, and near-miss events, so that etiologies can be discussed and hospitals can improve their service quality. Many parame-ters that require improvement through HTA (drugs, devices, systems) may not be discussed in this book.

Hospital-based HTA

HTA can and has even been suggested to be done locally at the hospital level, especially for important topics not assessed at the national level. In the large hospital third level of referral system, for instance, the effect-tiveness of a device vs. surgery to treat a certain disease should be evalu-ated. The assessment should include the implications of the methods, such as economic feasibility, etc. A complex technology may not be considered in a lower rank hospital, because it is not a priority for HTA nationwide.

Lower rank hospitals may have problems or topics less specific to larger hospitals. In a situation where a national HTA is lacking, hospitals have the right to conduct their own HTA. For example, an HTA in a limited form, could be done to assess whether a hospital needs to purchase a certain device.

Hospital-based HTA has been done for more than 20 years, but few reports mention the use of HTA results by hospitals themselves. One of the reasons is that hospitals rarely submit these types of reports for publication.

There are 4 models for hospital-based HTA according to international HTA standards (see **Figure I-8**):

1   **Ambassador model**: In this model, clinicians act as the ambassador of HTA in practice. They may not play a role in HTA, but they are the key persons for HTA in the hospitals.

2   **Mini HTA:** In this model, professionals play a role in the HTA process, by collecting data and giving input to the policy makers.

3   **Internal committee**: In this model, a committee consisting of a multidisciplinary group and representing various fields is established and held accountable to make assessments and provide recommendations for hospitals.

4   **HTA unit**: This model is a formal organization that works full time. The HTA unit is the highest level HTA organization in the hospital.

Hospitals are well-known as institutions with intensive modalities, full of professions and solid technology. Hospitals are places where technology can be applied to confirm diagnoses, to treat, and restore health. Law number 40 year 2009, article 5 on hospitals stated that the main task and function of a hospital is a) treatment service and recovery, b) maintenance and health improvement, c) education and training of human resources, and e) research and development as well as health technology assessment to improve health service and considering health science. Therefore, quality control and cost control is necessary in every hospital, so the hospital is the most suitable institution for having an HTA unit. Nevertheless, the HTA unit as part of quality and cost control can also be implemented in other institutions, such as the National Agency of Drug and Food Control (BPOM), the health district office, and BPJS Health.

	Focal activities	
	Clinical practice	Management
Low (Individual)	Ambassador model	Mini-HTA model
High (Team or Unit)	Internal committee model	HTA unit model

Figure I-8. Scheme of hospital-based HTA, from the most conservative (ambassador model) to the most comprehensive (HTA unit).

References

1   Brockis E, Marsden G, Cole A, Devlin N. A review of NICE methods across health technology assessment programmes: differences, justifications and implications. Office of Health Economics – National Clinical Guidelines Centre, Royal College of Physicians. London 2016.

2   Chaikledkaew U, Kittrongsiri K. Guidelines for health technology assessment in Thailand. 2nd ed. The development process. J Med Assoc Thai 2014; 97 (Suppl. 5): S4-S9.

3   Eddy D. Health technology assessment and evidence-based medicine: What are we talking about? Value in Health 12 (Supplement 2), 56-7, 2009.

4   European Coordination Committee of the Radiological, Electro- medical and Healthcare in Industry. COCIR position paper: Assessing the value of Medical Imaging and Health ICT The role of Health Technology Assessment (HTA). October 2014.

5   Goodman CS. HTA 101: Introduction to health technology assessment. Bethesda, MD: National Library of Medicine (US); 2014.

6   Guidelines for stakeholder engagement in health technology assessment in Ireland, 2014. Health Information & Quality Authority, 2014.

7   Health technology assessment of medical devices. WHO Medical Device Technical Series. 2011.

8   Hailey D. Local health technology assessment: A guide for health authorities. Alberta Heritage Foundation for Medical Research. Suite 1500, 10104 – 103 Avenue Edmonton, Alberta, Canada T5J

9   Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. Designing clinical research. Philadelphia: Lippincott, 2007.

10 James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J. 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults. Report From the Panel Members Appointed to the Eighth Joint National Committee (JNC 8). JAMA. 2014;311:507-520.

11 Kristensen FB, Sigmund H, editors. Health technology assessment handbook. © Danish Centre for Health Technology Assessment, Nation-al Board of Health URL: http://www. dacehta.dk/ 2007

12 Lysdahl KB, Oortwijn W, van der Wilt GJ, Refolo P, Sacchini D, Mozy-gemba K, et al. Ethical analysis in HTA of complex health interventions. BMC Medical Ethics 2016; 17:16:2.

13 O'Donnel JC, Pham SV, Pashos CL, Miller DW, Smith MD. Health technology assessment: lessons learned from around the world—an overview. Value in Health. 2009; S1-S5.

14 Petherick ES, Villanueva EV, Dumville J, Bryan EJ, Dharmage S. An eva-luation of methods used in health technology assessments produced for the Medical Services Advisory Committee. Med J Aust 2007; 187;5:289-92.

15 Ruiz F. Overview of "health technology assessment" and its role in decision making ARCH Initiative HTA Workshop. 2014

16 Sastroasmoro S, Ismael S, editors. Dasar-dasar metodologi penelitian klinis. (Principles of clinical research). 5th. Ed. Jakarta: Sagung Seto, 2014.

17 Sastroasmoro S, editor. Menelusur asas dan kaidah *evidence-based medicine.* (Reviewing the principles of evidence-based medicine). Jakarta: Sagung Seto, 2014.

18 Whyte P, Hall C. The role of health technology assessment in medicine pricing and reimbursement. WHO, http://haiweb.org/wp-content/uploads/2015/08/HTA-final-Aug2013a1.pdf

Part Two

Guidelines on Economic Evaluation and Budget Impact Analysis for Health Technology Assessment (HTA) in Indonesia

LIST OF ABBREVIATIONS

ACER		average cost-effectiveness ratio
BIA	:	budget impact analysis
BPJS	:	Badan Pengelola Jaminan Sosial
CBA	:	cost-benefit analysis
CEA	:	cost-effectiveness analysis
CEAC	:	cost-effectiveness acceptability curve
CHEERS	:	Consolidated Health Economics Evaluation Reporting Standards
CMA	:	cost-minimization analysis
CUA	:	cost-utility analysis
DALY	:	disability-adjusted life year
DSA	:	deterministic sensitivity analysis
HTA	:	health technology assessment
ICER	:	incremental cost effectiveness ratio
JKN	:	jaminan kesehatan nasional
PAH	:	pulmonary arterial hypertension
PICO	:	population, intervention, comparison, outcome
PRO	:	patient reported outcome
PSA	:	probabilistic sensitivity analysis
PTK	:	penilaian teknologi kesehatan
QALY	:	quality-adjusted life year
RCT	:	randomized controlled trial
RR	:	relative risk
SA	:	sensitivity analysis
VAS	:	visual analog scale
WTP	:	willingness to pay

Chapter 9

Introduction

Health technology assessment (HTA) is policy analysis that is conducted systematically with a multidisciplinary approach in order to assess the impact and use of health technology. The HTA process includes clinical, epidemiological, socio-cultural, ethico-legal, and economic parameters.

HTA in Indonesia was mandated by Government Regulation No. 12 year 2013 article 43, which states that in order to control the quality and cost of health care, the Ministry of Health is responsible for implementing HTA. The HTA Committee was established by the Ministry of Health, Republic of Indonesia, Decree No. 171/Menkes/SK/IV/2014, followed by Ministry of Health, Republic of Indonesia Decree No. HK.02.02/MENKES/422/ 2016. The Ministry of Health has developed plans for HTA implementation, with systematic mechanisms and credible institutions. One of the important components of HTA is an economic evaluation. This evaluation helps policymakers decide which health technologies to include in the benefit package of the National Healthcare Security (*Jaminan Kesehatan Nasional* – JKN).

Why does HTA require an economic approach? The goal of an economic evaluation is to address the problem of scarce resources for meeting unlimited human needs. Hence, methods and measures for assessing the worth of using a certain resource for achieving a certain result are needed. The principle of opportunity cost (often referred to as "benefit forgone") applies to the allocation of finite resources, as any decision to use these limited resources eliminates the opportunity for another stakeholder to use them. Thus, any decision to use the resources should be the "best" according to an economic viewpoint, which requires efficiency in resource allocation. Economic evaluation in the field of health care is done with the aim of discussing how the decision to use an intervention (curative services as well

as a public health approach) is done systematically, using accurate and credible information and specific methods of economic analysis. The appli-cation of economic evaluation in the field of health technology, particularly for drug interventions, is known as pharmacoeconomics.

HTA activity may be divided into two primary processes. The first process is the **assessment** to obtain evidence that a specific intervention has econo-mic value (referred to as a process to generate evidence or conduct an economic evaluation). The second process is the **appraisal** of whether the assessment was done according to standard. The first process is done by an "HTA agency," such as universities or research institutes, while the second process is carried out by the HTA Committee, which includes several experts. After both processes are completed, then a decision-making pro-cess is done to decide if the new intervention should be included or remain in the benefits package (or not).

Economic evaluation in HTA helps policymakers decide which health technologies (drugs, procedures, medical devices, and much more) have **value for money,** in order to be included in the benefits package. Moreover, economic evaluation can be also used as the basis to exclude an existing technology from the benefit package (disinvestment).

In terms of the continuum process, HTA begins with an assessment of whether a health technology is safe, effective, and efficacious, followed by an economic evaluation (cost-effectiveness or cost-utility analysis). The final step of the process is an analysis of the ability to fund these inter-ventions (affordability) or budget impact analysis (BIA) or financial impact.

Objective of the economic guideline

In general, this guideline aims to assist the stakeholders in assessing and conducting economic evaluations in the health sector, with the following expectations to be achieved:

- Consistency in HTA assessment (standard methods and report-ing the results of the study)

- Transparency and systematic results (systematic approach)

Scope

This economic evaluation guideline is intended to be a reference or stan-dard to conduct an economic evaluation as part of the HTA process in Indonesia. This guideline also includes an explanation of how to report the results of a good economic evaluation; however, the methods for appraising the results of an economic evaluation are separate from this guide.

HTA covers a wide range of health technologies such as drugs, devices, diagnostic and therapeutic procedures that are usually used in clinical settings, as well as screening, immunization programs, and other technolo-gies used in public health sectors. Pharmacoeconomic studies for drugs have grown rapidly, while studies in other sectors, such as for medical devices and public health programs, have not kept a similar pace. The examples used in this guideline are primarily for economic evaluation of drugs/pharmacological interventions.

The description of the implications for the budget or budget impact analysis (BIA) is used to help the payers (BPJS and the Ministry of Health) estimate the implications of how much money is needed for new interventions/ health technologies which have been proposed to the decision-makers, compared to the cost of the current intervention.

Chapter 10

Basic Concepts of Economic Evaluation

Definition

Economic evaluation, often referred to as the evaluation of economic efficiency, provides important information for policymakers, and answers the question of whether a procedure, service, or program is worth doing compared to the alternative, in terms of using limited resources.

Economic evaluation in the context of HTA is important, especially when paired with three other types of evaluations that answer different questions:

- Can this intervention work? This evaluation is related to efficacy.

- Does this intervention work? This type of evaluation relates to the effectiveness or benefit from the results. The fact is, results of efficacy trials, which are conducted in ideal circumstances and, tightly controlled, are quite different from success in the real world. Efficacy that is proven in an RCT (randomized controlled trial) is not necessarily equal to the effectiveness of treatment in the hospital. This question is often connected with "does it work in reality?" Is it worth sacrificing limited resources when it is successfully applied in the real world? The question is valid be-cause many factors affect the success of an intervention. The evaluation includes costs and consequences or outcome (cost-effectiveness analysis).

- Can this intervention be applied to and reach the people who need it? This issue is related to availability, as well as fairness and equity issues.

Why is economic evaluation important?

Reasons for the importance of conducting economic evaluations:

1 Without a systematic analysis, it is difficult to clearly identify the relevant alternatives to be compared.

2 To avoid bias in points of view (perspectives). Various parties in HTA enter the process with different assumptions. One alterna-tive may not look attractive using one assumption, but could prove better using a different assumption. As such, it is impor-tant to embrace many points of view, including those of the patients, the target group of a service, the budget from the Ministry of Health, other sectors, or society at large.

3 To help identify and interpret uncertainties that arise

Types of economic evaluation

Economic evaluation has two important and basic characteristics:

1 Input and output, often referred to as the costs and conse-quences of an activity

2 Selection. We are always faced with limited resources. It is impossible for an intervention program to fulfill all the needs of all the people. Although medicines or therapy may be highly effective, a decision is always required as to whether the medi-cine/therapy can be provided to society.

Thus, economic evaluation can be considered as "comparative analysis of two or more alternative interventions, both in terms of cost and conse-quences (specifically the results or outcome as a result of the intervention)." As such, the economic evaluation should be able to answer the question of, "compared to what?"

Comparisons may be made to the alternative intervention program in terms of costs and results; if no other intervention exists, the alternative is often called "do nothing." For example, one might compare the success of influ-

enza vaccinations for the elderly compared to no vaccinations at all. Is it worth the sacrifice (cost) compared to the outcome? For HTA done with the aim of including the intervention or health technology in the benefit pack-age, that health technology is compared to interventions that are currently available for the same purpose. See **Table II-1.**

Table II-1. Types of economic evaluations in the health sector

	If the study considers only the cost or consequence/outcome /output	If the study considers both cost and consequence/ outcome/output
No comparison to other alternatives	(1) PARTIAL ECONOMIC EVALUATION: Covers only one, either (a) Outcome description, or (b) Cost description	(2) PARTIAL ECONOMIC EVALUATION: As a cost-outcome description
Comparison between two or more alternatives	(3) PARTIAL ECONOMIC EVALUATION: Covers only one, either (a) Evaluation of the efficacy or effectiveness, or (b) Cost analysis	(4) FULL ECONOMIC EVALUATION: - Cost-minimization analysis (CMA) - Cost-effectiveness analysis (CEA) - Cost-utility analysis (CUA) - Cost-benefit analysis (CBA)

Source: Drummond et al, Methods for the economic evaluation of health care programs, 2015.

The study which evaluates an intervention program (health technology), without a comparison as shown in **cell 1** on **Table II-1**, yields merely an outcome description or cost description. Economic studies of "cost of illness" or "economic burden of disease" are included in this group, but they are considered to be partial, without a comparison. An analysis to calculate the unit cost of service (outpatient or inpatient) is also included as

an example in cell 1. Similarly, we might compare the unit cost of services at two hospitals or health centers; these examples are the results of the cost analysis (cost per output unit regardless of the outcome in effectiveness).

In **cell 2**, cost and outcome of an intervention are measured, but with no comparison. This evaluation is referred to as a "description of the costs and outcomes." For example, a study was done by Reynell and Reynell in 1972 on service in a coronary heart disease unit. The results described the cost of intervention in the coronary heart unit of a hospital and estimated the number of lives saved. The authors did not compare the intervention to other alternatives. As such, though the researchers refer to this study as a cost-benefit analysis (CBA), according to the rules for economic evaluation, a study without a comparison is not a full economic evaluation.

In **cell 3**, two alternatives are compared, in terms of cost and outcome, but they are not examined simultaneously. For example, only the efficacy (evaluation of efficacy or effectiveness) or only the costs of the two alternatives are compared. This evaluation is referred to as a study of the efficacy or effectiveness or a cost-only analysis. Efficacy trials to compare a new medicine to the standard medicine are included in this category.

Lawson *et al.* examined the cost differences of three methods of in-house, long-term oxygen therapy, such as oxygen cylinders, liquid oxygen, and oxygen concentrators/machines to extract oxygen from the air. The researchers argued that a cost-analysis would be adequate for this study because the effectiveness of the 3 methods was not a substantial problem.

Although this study is included in the category of partial economic evaluation, it does not mean that it was not important. Studies such as these can quickly provide information about the costs and outcomes of an intervention to the stakeholders. For example, we might use a partial economic evaluation to advocate to policymakers on economic loss, when a particular disease is not handled properly due to lack of attention or limited resources. Studies to answer or explain the issue of economic efficiency require a full economic evaluation, measurable in scope, and carried out systematically so the cost and the success of an intervention can be seen compared to the standard.

In **cell 4**, the category of "full economic evaluation" is done to compare two or more interventions, both in terms of cost and outcomes as well as the link between the costs incurred compared to the outcomes.

The principle of full economic evaluation is to explain economic efficiency, in which the sacrifice of resources should be in keeping with the outcomes obtained. Some interventions require greater costs, but achieve greater results. In contrast, some interventions are more efficient than slightly more costly interventions, but have unsatisfactory outcomes.

Included in this category are the **cost-minimization analysis (CMA), cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost-benefit analysis (CBA)**. CEA and CUA are the most commonly used in HTA economic evaluations. CMA and CBA are also done in several other coun-tries, with regards to certain conditions and the need for policy input. See **Table II-2**.

Cost-minimization analysis (CMA)

CMA is the simplest type of economic analysis for health technology. CMA is used to compare two or more health technologies that have similar, equivalent, or deemed to be equivalent clinical outcomes. Since the out-come of two or more of these interventions are considered to be the same, then a comparison need be made of only the costs.

For example, one might compare the costs of surgery with hospitalization vs. surgery without hospitalization (patient is allowed to go home after the operation is completed). If the clinical outcome according to the surgeon is similar, then only the costs incurred for both types of patient management are compared.

In general, CMA was not planned from the start for economic analysis, because it was considered to be too modest to simply compare the cost of which intervention was "cheaper" (e.g., the cost of medicine or treatment). However, when the clinical outcomes are the same for two types of inter-ventions, the researchers can just compare the costs of the two interven-tions.

Table II-2. Types and characteristics of economic evaluation

	Numerator	Denominator
Cost-minimization analysis	IDR (Indonesian Rupiah)	-
Cost-effectiveness analysis	IDR	Health process or outcome in natural unit (e.g., mmHg or years of life gained)
Cost-utility analysis	IDR	Outcome in unit form (e.g., QALY, DALY)
Cost-benefit analysis	IDR	IDR

Cost-effectiveness analysis (CEA)

CEA is a comparison of two or more medical interventions that provide different magnitudes of outcomes. As this analysis simultaneously measures the costs and the effects, users can determine which medical intervention is most efficient for obtaining the desired results. For example, a study might compare two or more kinds of medicines from the same thera-peutic class, but that provide different magnitudes of treatment outcome, such as two anti-hypertensive drugs that decrease diastolic blood pressure differently.

In CEA, intervention costs are measured in monetary units (IDR) and outcomes are measured in units that correspond to the clinical outcome or health indicator, e.g., mmHg decrease of diastolic blood pressure, number of cataract surgery cases with certain costs (different procedures), or the number of deaths that can be prevented or life years saved (e.g., breast cancer screening programs, vaccinations, and other preventive efforts).

CEA is used to compare medical interventions that have the same purpose, generally between a new intervention and a standard of comparison. The CEA generates two ratios that provide different information: the ACER (average cost-effectiveness ratio) and ICER (incremental cost-effectiveness ratio). The ACER directly compares the ratio of costs and effectiveness of a single intervention whereas ICER is the ratio of the difference between the costs and the difference between the outcomes (DALY or QALY) of the interventions.

The final result of CEA in HTA is the ICER. The ICER is compared to a threshold value to assess if the intervention has value for money; i.e., if the ICER is below the specified threshold, then the intervention is considered to be cost-effective. Thus, in HTA the ACER only reveals the cost to benefit ratio of an intervention without reference to a comparison. However, ICER reveals if the cost difference between the new and old interventions is compatible with the increase of the obtained outcome (e.g., if drug A compared to drug B has value for money). Then, ICER is compared to the specified threshold.

$$\text{ICER} = \frac{\text{Cost of new intervention} - \text{cost of previous intervention}}{\text{New intervention outcome} - \text{previous intervention outcome}}$$

Cost-utility analysis (CUA)

CUA is similar to CEA, but the outcomes are expressed as the utility associated with changes in length of life and quality of life, as a result of a health intervention. Outcomes in the form of length and quality of life reflect the state of the following:

- Does the illness shorten the life of the patient and/or reduce the patient's quality of life?

- Can medical interventions improve the patient's length and/or quality of life?

A standard of outcome in CUA is quality of life, developed from the concept of utility, or the patient's level of satisfaction after receiving a medical

service, e.g., after treatment for cancer or heart disease. The unit of utility is calculated as 'number of years adjusted' or quality-adjusted life years (QALY), after considering the additional of life years of these patients. QALY is measured using two approaches: quantity (length of life) and quality (quality of life). Further explanation can be found in sub-part of the outcome (Chapter 11).

As in CEA, the final result of CUA is the ICER value, with the numerator being the difference in costs of the new and previous interventions (drug), and the denominator being the difference in the values of effectiveness/ outcome of the new and previous interventions (drug), in the form of QoL.

$$\text{ICER} = \frac{\text{Cost of new intervention} - \text{cost of previous intervention}}{\text{QALY new intervention} - \text{QALY previous intervention}}$$

The ICER result is the cost/QALY gained (cost to acquire one additional year of quality life). This value is then compared to a pre-determined threshold value to decide whether it has "value for money" (in this case, a cost-effective intervention would have a value is below the threshold defined by the state).

Cost-benefit analysis (CBA)

CBA is applied to comparisons of two or more interventions with different objectives and outcomes. Both costs and outcomes are measured in mone-tary value (IDR), and adapted to the period of calculation (discounted). The outcome of CBA is a benefit, measured in units of currency, as the cost saved if the analyzed interventions are obtained/successful. Alternatively, CBA can be a way to measure the willingness to pay (WTP).

The basis of CBA is a surplus of benefits (net benefit), i.e., the obtained benefit minus the cost. CBA can also be used to compare the benefit ratio with costs. If the value is positive, then the intervention has "value for money," and can be accepted for implementation. The higher the CBA value, the greater the value-added, thus increasing the likelihood of accept-ing the newly-proposed intervention. For example, the benefit of a success-

ful malaria eradication program is cost-saving, as it comprises the cost of outpatient and inpatient treatment, medicines, productivity loss, and so on. This guideline is intended to focus on CEA and CUA of HTA, not explain CBA in detail with examples. Further explanation can be obtained from the references in this guideline.

Chapter 11

Steps in economic evaluation for HTA

Economic evaluation framework for HTA

Economic evaluation as part of HTA must have a clear framework as the basis of the health technology selection process. Types of economic evalua-tion and their interpretation provide input and guidance for the decision-makers (evidence-informed decision making), especially to determine if certain health technology will be included or excluded from the benefit package of BPJS.

For that reason, we need to understand the following:

- Selection of the technology topic to be studied, such as drugs, devices, medical or surgical procedures, or other health technol-ogies.

- Development of research question(s).

- Assessing the background of the health technology to be studied, including reviewing relevant literature, the context of the devel-opment and needs of our country, as well as available alterna-tives and their consequences.

- Outcomes to be measured, such as efficacy and effectiveness, the procedure itself (e.g., the treatment), resources used (cost), as well as legal, social, religious, and ethical aspects, must be clearly described.

- Data analysis (statistical or modeling) and their interpretation and may also have requirements for implementation.

Steps in economic evaluation

Step 1: Determine the direction, scope, and study protocol

The initial step in the economic evaluation for HTA is to establish the assessment framework. Select the topics to be studied, which have been chosen by systematic procedures of HTA (see other documents regarding HTA institutionalization). The incoming proposals to be followed up with an economic evaluation are usually accompanied by a standard of compari-son, or to be determined by the researchers on the basis of a decision from the HTA Committee. The establishment of study protocols includes the following important aspects explained under Step 1.

Literature review

Key points
a. All relevant articles should be critically appraised using standard methods, and combined systematically with credibility and transparency.
b. The systematic review (SR) results help researchers to develop theoretical and conceptual frameworks for economic evaluation through modeling
c. The SR results are needed to obtain the parameter values to be used in modeling in the economic evaluation

In economic evaluation, critical review of the published articles in health outcomes and health economics is very useful for giving an initial picture, before conducting the actual health technology assessment. The literature review is also beneficial for providing relevant information for the econo-mic evaluation model development plans and parameters. The literature review should be carried out properly and critically, as it is not only a summary of methods and results of each study.

A systematic review (SR) is a structured and transparent procedure to summarize all available evidence from various studies. The SR should include the following steps. **First**, inclusion and exclusion criteria must clearly defined, specific, and relevant to the study. The PICO (population, intervention, comparison, and outcome) approach is recommended to facilitate researchers in the literature review. **Second**, the search strategy should be systematic and reproducible, by using relevant databases and appropriate terms. The use of large databases such as EMBASE, MEDLINE, PubMed, Cochrane, and many more are highly recommended. Studies done in Indonesia, but not available in international journals should also be included. **Third**, the articles should be managed by referring to the inclu-sion criteria. Avoid duplication of articles.

Fourth, summarize all findings from the extracted articles by reading the full texts, followed by critical assessment.

Research questions

Key points
a. Formulate a structured study protocol.
b. Specified and well-defined research questions are the important first step.
c. Information related to the disease and the type of economic evaluation should be clearly stated.
d. Health technologies (medicines, medical devices, or others) assessed should be mentioned.

The purpose of a study protocol for economic evaluation is to answer a question. We have to clearly specify the research questions and objectives of the study. What interventions are to be valued? To what interventions will they be compared? Again, it is important to remember the "PICO approach":

- Target population (Population=P)

- Technology or intervention (Intervention=I)

- Comparison (C)

- Expected outcomes to be achieved (Outcome=O)

The SR process at the beginning can provide an overview of the effectiveness assessment obtained in previous studies, the submitted research questions, and the most common and relevant methods used. The results will help researchers to formulate the research questions. By performing a SR, we can gain a clear understanding about the effectiveness of the technology in previous studies, the research questions to be formulated, and methods used. The results of the SR can be used by researchers to develop clear and focused questions, as a basis for conducting research.

The research questions should be clear, realistic, answerable, and in accord-ance with the desired context by decision-makers. More specific questions, secondary questions, and sub-group level questions can also be developed as required. For example, are additional analyses with more specific research questions required for stakeholders' understanding?

Information about the disease, the type of intervention, and economic evaluation should be mentioned. The name of the health technology (medicine, medical device, etc.) should be mentioned, accompanied by necessary explanations, e.g., brands, dose, and type (oral or other).

The target population

Key points
a. Determine the target population and benefits in order to develop the appropriate assessment.
b. Inclusion and exclusion criteria should be clearly defined.
c. Explain the process of data collection and analysis for a transparent and credible HTA.

The target population is defined by age, gender, socioeconomic status, type of disease, type of intervention, and other factors. For example, if we wish to perform a HTA on sildenafil for patients with pulmonary arterial hypertension (PAH), then, in this case, the patients are the PAH population. The inclusion of PAH patients should be clear. Will adults, children, or both be included? Then, a reason for performing the HTA on sildenafil for PAH should be given, i.e., PAH is not covered in the benefit package, hence, the need for a HTA study.

Data may have been collected from hospitals already using sildenafil for PAH treatment. If so, then we should mention the reasons for selecting these hospitals, followed by a description of how the data were collected from selected hospitals (sources, samples, etc.). To support the findings, sub-group analyses or analyses related to the heterogeneity of the target population may be performed, if necessary.

Comparison

Key points
a. The interventions/health technology comparison should be clearly described
b. Standards of comparison are determined based on interventions which are most often used to treat a disease, before any new technology was available for assessment

The comparison must be in accordance with the clinical context of the studied cases and scientifically proven. Selection of the comparisons must be accompanied by reasons as to why X is more relevant than Y, and why other options are considered to be less relevant. Comparisons are determined based on the most commonly used intervention before the new health technology was introduced, or the comparison may currently be in the guaranteed benefit package.

For example, we want to assess if a new technology is cost-effective for patients with acute myocardial infarction (AMI). The proposed technology is thrombolytic therapy with tissue plasminogen. The previous intervention, which is fairly common for AMI patients, is streptokinase. As such, streptokinase would be the standard of comparison in the study. The chosen comparison(s) are usually the routine standard of usual care. Let's take another example. In the case of terminal renal failure, we do not compare hemodialysis (HD) directly to peritoneal dialysis (PD). Each one, HD or PD, should be compared to treatment without dialysis at all. The reason for not performing a "head-to-head" analysis of HD and PD, in addition to scientific evidence in support of policy, is that both HD and PD will remain covered in the JKN. In a study on the proposal to extend the coverage of PD in the benefits package, the economic value of the PD package should be assessed to decide if it has a good value for money.

Choosing the type of economic evaluation in HTA

Key points
a. The recommended evaluation for HTA is the CUA, which can be compared to the threshold value.
b. Economic evaluation studies are carried out according to the steps in the cost-effectiveness analysis of CUA.
c. HTA in Indonesia uses both societal and provider perspectives.

Economic evaluation methods used in HTA are the CEA and CUA. Economic evaluation of new health technologies/medicines includes an analysis of the cost of the intervention and the outcome (often called "effect" or "consequence"). Economic evaluations utilize primary data to obtain patient-level clinical outcomes of an intervention or new health technology, or by doing a CEA alongside a clinical trial/observational study. An economic evaluation can also be obtained from the modeling results, referred to as "model-based economic evaluation" studies.

The differences in costs and outcomes of two interventions should be compared to each other, as to which are dominant, dominated, or a trade-off. The cost-effectiveness (CE) plane diagram explains the positions. See **Figure II-1**.

CE PLANE
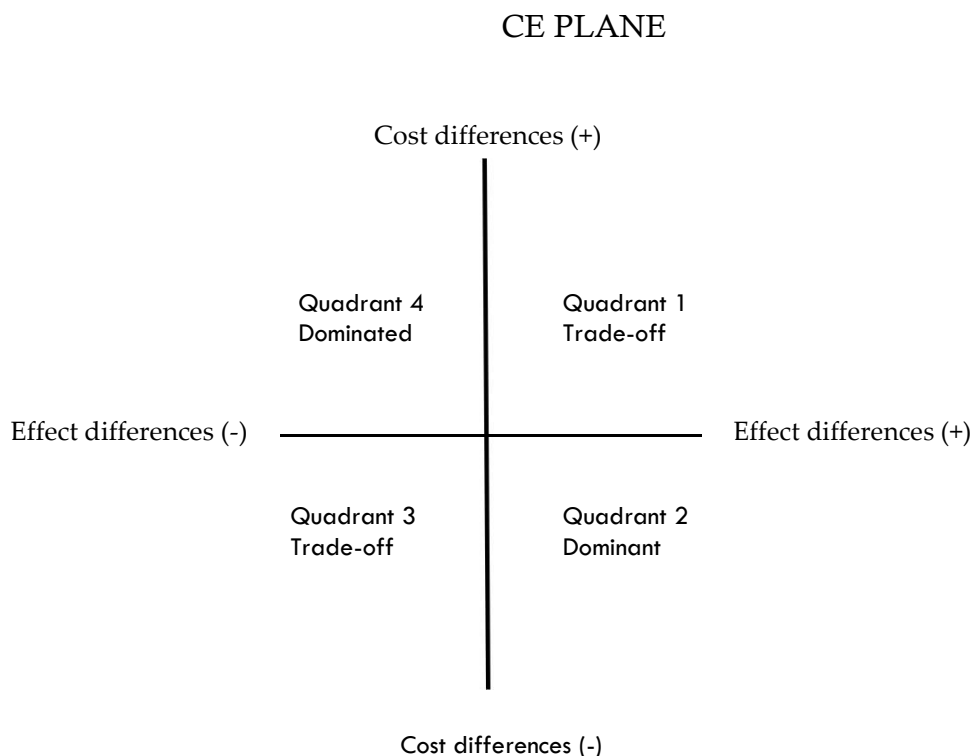
Figure II-1. Cost-effectiveness plane

In **quadrant 1** of **Figure II-1**, the new intervention has superior outcomes to the previous intervention, but it also has higher costs. The situation is said to have a "trade-off" (the exchange balance between the outcome at a cost). In this case, the proposed new intervention requires "value for money," that is, the higher cost of the proposed intervention must be compatible with the increase in outcome.

In **quadrant 2**, the new intervention is dominant to the current intervention (cost of the new intervention is lower than that of the current intervention,

while the outcomes are higher). In these circumstances, it is clear that the proposed new intervention is superior to the current intervention.

In **quadrant 3**, the required cost for new intervention is lower than that of the standard of comparison, but the outcome is not better (slightly lower or equal). In this case, the new intervention must be proven acceptable through a study.

In **quadrant 4**, the new intervention is dominated by the standard of comparison. Hence, the new intervention would be rejected. The CE plane graph can be obtained from a sensitivity analysis.

The steps according to the selected type of economic evaluation

The CUA is a type of economic study recommended for HTA. In accordance with protocol, CUA steps include cost and outcome data collection and analysis. Clinical outcomes data are obtained from the SR and meta-analysis results, while data on cost and utility are derived from primary data in selected locations. Effectiveness is given in units of QALY (quality-adjusted life year) or DALY (disability-adjusted life year).

1. After defining a clear policy question and the PICO parameters, one should explore various references related to intervention outcomes by SR and meta-analysis. The results will provide information about similar studies that have been done. The SR provides both the reference to build the framework of the modeling concept and the results of parameter data used for modeling.

2. Develop an appropriate modeling design for the planned analysis (decision tree or Markov model). Complete the required parameters: cost and clinical outcomes (relative risk, risk ratio, or others). Build a model involving the relevant stakeholders, including clinicians related to the study.

3. Collect data on costs (from societal and provider perspectives) and utilities, according to the number of subjects and instruments that are being developed, both in the intervention and the com-parison groups. Perform appropriate modeling

with a prede-fined time horizon, with discounts for selected parameters, as well as survival analysis.

4   Perform an analysis to generate adjusted cost and outcomes, including life years gained and QALY results.

5   Calculate incremental cost, incremental effectiveness (the differ-ence between the cost and effectiveness/intervention outcomes with the comparison), and ICER.

6   Perform a sensitivity analysis.

7   Interpret the ICER results and compare to the threshold.

The final result of the HTA study can be either cost per QALY gained or cost per DALY averted, and compared to the threshold (commonly referred to as willingness to pay or WTP), which has been established by the government or policy makers.

The perspective used in economic evaluation

The perspective on economic evaluation should be specified in the proposal/study protocol as well as in the final report of the study results. The types of perspectives in economic evaluation based on the cost incurred are:

- Societal

- Payer

- Patient

Economic evaluation is used to assess the efficiency of alternative health interventions, with an economic approach concentrating on society's welfare. The results of an economic evaluation will impact the whole society, not only organizations or individuals. Therefore, HTA in Indonesia is expected to use a societal perspective. The perspective must be expressed either in the protocol or the final report. As such, the resource and cost assessment are done from a societal perspective (for economic evaluation), and a provider perspective (for economic evaluation and budget impact analysis). The collected data must be presented in accordance with the planned perspective.

Step 2: Data collection

Data is crucial to economic evaluation. First, data related to effectiveness should be well specified. Ideally, if available, local primary data from our own country would be used. However, clinical studies in Indonesia are considerably limited, and would require more time and expense to conduct a special-focus study related to the technology assessment. Moreover, evidence obtained from only one study may be less useful for HTA. Most other countries apply the systematic review for data collection.

In terms of cost and utility, HTA Indonesia recommends using primary data collected in Indonesia. Primary data collection needs to be vetted thoroughly by an ethical clearance process. Data and information for economic evaluation include: (1) Efficacy, safety, and effectiveness, (2) Time horizon, (3) Cost, (4) Type of outcome, and (5) Discounting.

Efficacy, safety, and effectiveness

Key points
a. Gather data related to efficacy, safety, and effectiveness
b. SR results should provide a strong level of evidence
c. If the resources or SR are limited, an explanation will be required. As such, the limitations may impact the model's result

RCTs to assess an intervention and a standard of comparison have a high level of internal validity and generate evidence. However, regardless of the results obtained, the conditions in a strictly controlled study are not exactly the same as conditions in the real world. As such, issues of efficacy and effectiveness should be described. In order to register a drug at the BPOM, quality, in addition to safety and efficacy, must be considered. For economic evaluation, effectiveness is important because it describes real world condi-tions. Data related to efficacy are generally available, while data related to effectiveness and cost are usually collected from primary data corres-ponding to the country's conditions.

A SR is used to obtain clinical outcome data. If possible, one should conduct a meta-analysis to obtain a pooled estimate as a comparison to the inter-vention/medication. The limitations of the SR or the reason for not conduct-ing a SR should be clearly explained, such as due to limited number of publication sources or weaker levels of evidence.

The result of the SR is very dependent on the availability of information or publications. For further information, one can read The Handbook of the Cochrane Review. The steps for conducting SRs are explained in another section of this guideline (**Chapters 4** and **5**).

Time horizon

Key points
a. Specify the time horizon
b. The time horizon should be similar both for cost and outcome
c. Use a long time horizon, supported by justifications

It is important to note that time horizon applied to the analysis should be long enough to accurately capture the cost and outcome consequences. A less than appropriate time horizon may lead to bias in the decision analysis process. Primary data or data from a RCT with a short time horizon, could potentially be extrapolated to a longer period beyond that clinical trial (by means of modeling).

Cost and outcome are frequently shown in different time periods. For instance, in an economic analysis of immunization programs, costs are incurred in the present time, while the benefit can be perceived as a lifetime. This type of case utilizes money now and delays the outcome, and is related to the 'time preference concept.' Hence, a long time period should consider a discounting factor (discount rate) for adjustment, both in cost and the outcome.

Cost

Key points
a. For HTA Indonesia, the economic evaluation uses a societal perspective
b. Cost identification and calculation must be conducted in a systematic and transparent way, using primary data
c. Adjust future costs to present values by discounting

Cost is a resource sacrificed for the purpose of achieving a certain outcome, measured in monetary terms. Cost not only consists of reported costs and transaction results, but has an expanded definition that includes the con-cept of opportunity cost. Opportunity cost is a loss of the chance to use a resource elsewhere due to its utilization for other activities. Cost also can be viewed from different perspectives.

Cost according to a hospital perspective is a cost required to provide health services/care for patients, e.g., to provide outpatient and inpatient services in a hospital. In general, a cost analysis generates actual (historical) cost. Hospital costs can be categorized as follows:

- Fixed cost and variable cost. Fixed cost is a cost not influenced by output volume, while variable cost is influenced by output volume. For example, the cost of building a clinic for outpatient services is a fixed cost, but the cost of drugs/reagents is a variable cost.

- Investment cost and cost of operational maintenance

- Cost-related building in an investment context that assumes lifetime utilization should be taken into consideration, because the investment is used for a long term period. As such, the cost estimation should be annualized; this action is familiarly known as the annualized investment cost. The operational cost of main-tenance is the cost to run a corporation, in this case, a hospital.

Operational costs include the cost of a drug supply, utilities (electricity, telephone, water), building maintenance, equipment, etc.

- Direct cost and indirect cost (overhead cost). Direct cost is a cost used directly to generate health care services, e.g., all costs of inpatients, outpatients, laboratory, surgical unit, etc. consisting of fixed and variable costs. Indirect cost is the cost of a support unit in the hospital that must be imposed on a production unit, e.g., the cost of a maintenance installation in a hospital, the kitchen, the laundry, etc). Systematic cost analysis is needed to obtain the specific costs in the hospital.

- Total cost and unit cost. Total cost is all costs incurred by the hospital to produce health services for patients. The total cost from each production unit should be calculated in order to obtain a cost per patient, after adjusting for the number of patients in a certain year. For example, the total outpatient cost in hospital X was IDR 10 billion in 2014, with 40,000 patient visits that year. Thus, the unit cost would be 10 billion / 40,000 = IDR 250,000 Similarly, outpatient cost/length of stay and laboratory cost per test are unit costs.

Cost according to a patient perspective is the cost of obtaining health services. The cost incurred to obtain healthcare services due to illness is known as the cost of treatment. These costs include direct medical costs, that is, the cost to obtain healthcare services such as drugs, physician services, laboratory tests, etc. Other costs include costs related to the effort of obtaining the services, such as transport and other indirect costs related to lost patient productivity due to illness and hospitalization. For example, a patient may lose income due to their own illness, or a patient's family may lose income while accompanying/caring for an ill family member.

If patients have BPJS health insurance, then the cost can be estimated by the Ina CBGs tariff payment. BPJS provides healthcare services for particular cases or health conditions. If patients are covered by another health insur-ance scheme or their workplace, then the reimbursement claims should be calculated accordingly.

Sometimes the insurance claims for the healthcare fee are not fully covered so the patient must pay the remainder fee, called the "out of pocket pay-ment." This cost is estimated by the billing data. In order to prevent under-estimation or double counting in calculating patient costs, a structured interview is needed to capture the complete description about cost compo-nents that should be calculated.

Cost according to a payer perspective is a cost incurred by an insurer, in this case, BPJS, in the form of a hospital claim for health services. The reimbursement to the hospital is determined by the Ina CBGs tariff for a particular diagnosis. For example, in the case of renal failure or heart disease, this payment is the responsibility of BPJS in a package (bundled) form. The hospital cannot bill the BPJS for specific items, such as drugs, medical services, etc.

Cost according to a societal perspective or societal cost is defined as capturing all cost components, including patient costs (direct medical cost, direct non-medical cost, and indirect cost).

In cost calculation, we should avoid double counting, which is calculating the same cost from a different perspective. For example, an analyst calcu-lates patient costs for medication in a polyclinic at hospital X (the charges paid by the patient) and calculates the unit cost per visit to the polyclinic at that hospital.

Ideally, the sample for cost data represents variations that exist in Indone-sia. For example, cost data related to transportation are indeed different, between eastern and western Indonesia, or between urban and rural areas. As a consequence of limited time and resources, it is not feasible to include all data across Indonesia, so the sampling justification should be explained clearly.

Modeling on the basis of Bayesian statistics is not intended to prove the hypothesis (prove the influence of one variable on another variable). Parameter values included into a model template are the values that will be analyzed in consideration of uncertain aspects, then explained further in a sensitivity analysis.

Types of outcome

Key points
a. The final form of a clinical outcome is life years gained and the final form of a non-clinical outcome is QALY/DALY
b. Data sources for clinical outcomes are obtained from a SR, gathered from Indonesian and international sources.
c. To obtain QALY, a utility measurement using EQ-5D is collected as primary data, according to an Indonesian context

Outcome measurement for economic evaluation

In economic evaluation, outcome or effectiveness is defined as clinical outcome and non-clinical outcome.

Clinical outcome

Clinical outcome is a measurement of the effect of a drug intervention/ procedure on disease progression and on health status improvement. There are several data sources for clinical outcomes, including research that used primary data or SR results, and precise meta-analyses (from credible public-cations). HTA Indonesia recommends that the outcome be final out-come, such as survival years.

Studies conducted to measure final outcomes need an appropriate epide-miological design and may be costly. Clinical outcome also can come from efficacy and effectiveness studies in the BPOM registration process. Good SRs and meta-analyses as references have been proven valid and reliable.

In terms of efficacy, the highest hierarchy of evidence comes from RCT studies. RCT subjects are chosen with specific inclusion and exclusion criteria, followed by randomization to decide who receives the new inter-vention or the placebo/standard intervention.

An economic study that uses efficacy data as the outcome is called a cost-efficacy analysis (CEA). For CEA, RCT results are of limited use, because

the time horizon is generally too short a period to explain the outcome, in contrast to a long term period to prove patients' survival. Furthermore, strict conditions or ideals in RCT studies influence patient behavior that may be different in the real world. Thus, an effectiveness study is needed to provide information that is more reflective of events in the real world/ environment. For example, differences in access to healthcare services, cost required, patient adherence, medical standards, or social factors in the real world may greatly influence the portrait of resource allocation.

Here are examples of search results on effectiveness/clinical

outcome. *Dialysis Study*

In 2016, a dialysis study was performed as an HTA analysis by the Ministry of Health. A thorough search of PubMed and Cochrane led to identification of 606 articles. After screening based on title and abstracts, 11 articles remained. Finally, only two articles were included in the final review, based on inclusion and exclusion criteria.

Reference	Study design	Level of evidence *
Vale L et al., 2004	Systematic review CAPD vs. hospital/home HD in adult GGT	1-
Korevaar et al., 2003	Analysis of initial therapy HD vs. PD	1-

Non-clinical outcome

Non-clinical outcome is a measurement of quality of life, as number of life years gained in patients who used a health technology. Currently, HTA in several countries recommend that a patient perspective be included in mea-surements of non-clinical outcomes. As such, non-clinical outcome utility is important. Non-clinical outcomes that have been developed based on utility also must be presented for CUA. The preference to use DALY or QALY depends on the study context. In general, studies related to drugs and including a patient perspective use QALY as the outcome.

DALY (Disability-adjusted life years)

Macro-studies related to disease burden tend to use DALY more frequently. Also, DALY is applied to economic evaluation studies with value cost/ DALY averted (or avoided), which means the cost to prevent losing healthy years of life. The formulas are presented below:

- DALY= YLL+YLD

- YLL= years of life lost

- YLD= years living with disability

- One DALY $\rightarrow$ one healthy life-year lost

QALY (Quality-adjusted life years)

A QALY measurement in economic evaluation is the cost/QALY gained, or cost for an additional one year of healthy life, based on an estimate using a utility approach.

Steps to calculate QALY:

1. Create the description of each disease state

2. Specify the appropriate methods to estimate the utility (for HTA we use EQ-5D)

3. Specify the subjects/respondents that will be measured for their utility-related health status from each stage of health conditions

4. The utility value is multiplied by the number of life years expected from each option for gaining QALY.

Utility

Utility is a quantitative expression for an individual preference to achieve a particular health state under uncertain conditions. Preference is a general concept for utility and value. Preference measured in uncertain conditions is called utility, while preference measured in certain conditions is called value. Sometimes, both utility and value are called utility.

Why do we need to measure utility? A utility score reflects the preference of health status and the effect of health status improvement, as an impact of efficacy or side effects from the health intervention. A utility score is used for weighting the quality of life value when estimating QALY; it is possible for the combination of values to change either for morbidity or mortality in one result, hence, the QALY.

QALY is a combination of measurements of survival and quality of life of patients, in relation to utilization of health technology. Determination of the threshold ICER as the result from CEA also depends on final QALY result. For example, in the UK, NICE set ICER as £30,000/QALY gained for reimbursement by NHS, while the US set ICER as $50,000/QALY gained.

The question of whom to include in the QALY measurements has not been clearly answered, due to different perceptions. Some experts argued that in addition to patients themselves, those who become caregivers should be respondents (e.g., parents of a pediatric patient, or family members caring for patients surviving with dementia). Others argued that even the health care provider or society can act as respondents. There has been no consen-sus related to this ethical concern.

Patient measurements are familiarly known as patient reported outcomes (PRO), as a general term that includes outcome data which may be directly reported by patients. The PRO is used to explain the patient conditions and the results of medication. For instance, patient information collected may be related to functional status, well-being, symptoms, HRQL, satisfaction, and adherence. Some instruments used to measure such outcomes are the patient-reported outcome, caregiver-reported outcome (for dependency or functional status), clinician-reported outcome (for global impression, observation, or test of function), and physiological outcome (i.e., HbA1C, tumor size).

Methods of utility measurement: preference and value

There are three frequently used ways to measure utility preference and value by rating: the visual analog scale (VAS), the standard gamble (SG), and the time-trade-off (TTO). The disease state or condition is scored be-

tween 0.0 (dead) and 1.0 (perfect health). The SG and TTO are considered to be direct measurements. The SG generates a utility score of preferences under uncertainty, while TTO and VAS are used for value under conditions of uncertainty. SG is a preference method for personal decisions that are based on the maximal utility principle expected from decision theory. See
Table II-3.

For example, a respondent has a new wheelchair. The question asked of the respondent is, "What mortality risk could you accept for your recovery compared to your life with a wheelchair?" If the answer is 20%, then utility X is 100-20%, that is 80%, or 0.8.

A rating scale is a simple way to express numeric values of health status, in the form of a thermometer with the number 100 (healthy) at the top, and the number 0 (dead) at the bottom.

A visual analog scale (VAS) is the easiest approach when providing choices on a range of numbers, to determine patient health status. A question that can be used is "How about your quality of life?" However, although this approach is easy, VAS has several limitations. One such limitation is that VAS is unable to describe an interval. Respondents who complete a VAS tend to perceive the scale as having category-ordinal values, so a VAS is potentially misleading. Let's take an example.

Patients are asked to a choose number on the VAS from 0 (death) to 100 (healthy/no functional symptom). If a patient chooses 70, the patient health status would be a score of 0.7. A patient with a mild disease condition would select a value near to 1 (or 100), while a patient with severe disease would select a value close to 0 or even below the available scale. Fur-thermore, the length of time that a patient has suffered from a disease also influences their perceptions when giving scores.

An alternative to a rating scale is the TTO approach. Respondents are asked to choose (trade) length of life for quality of life. Using the wheelchair example, the subject would be asked, "If your life expectancy is 50 years with a wheelchair, how many years would you give up to have a healthy life without a wheelchair?" If the subject's answer is 10 years, we can estimate his quality of life by the following formula:

- QALY (wheelchair) = QALY (cured)

- Length of life * Quality of life (wheelchair) =LL*QoL (cured)

- 50*quality of life=40*1

- QoL (wheelchair)=0.8

Generally, a rating scale questionnaire is self-administered. However, SG and TTO are better conducted by direct interview because subjects may need clarification.

Indirect measurements use a utility-weighted index. Some examples include the EuroQoL, Health Utility Index (HUI), Quality of Well-being scale, or WHO QoL instruments. A tool for measurement of multiple indexes is the Multi-Attribute Utility Instrument (MAU instrument).

Choosing the appropriate measurement instrument should be based on a specific intervention. However, for HTA in several countries, EQ-5D is a good choice. This is not to say that other instruments are not good. But the EQ5D has a number of strengths, such as being easy to obtain, general, easy to understand, and the results can be used to measure perceived quality of life for patients with different diseases.

Table II-3. Methods for measuring quality of life

Response Method	Question Framing	
	Certainty (Values)	Uncertainty (Utilities)
Scaling	1. -Rating scale -Category Scaling -Visual analogue scale -Ratio scale	2.
Choice	3. -Time Trade off -Paired Comparison Equivalence -Person trade-off	4. Standard Gamble

Source: Drummond et al, Methods for the Economic Evaluation of Health Care Programmes, 2015.
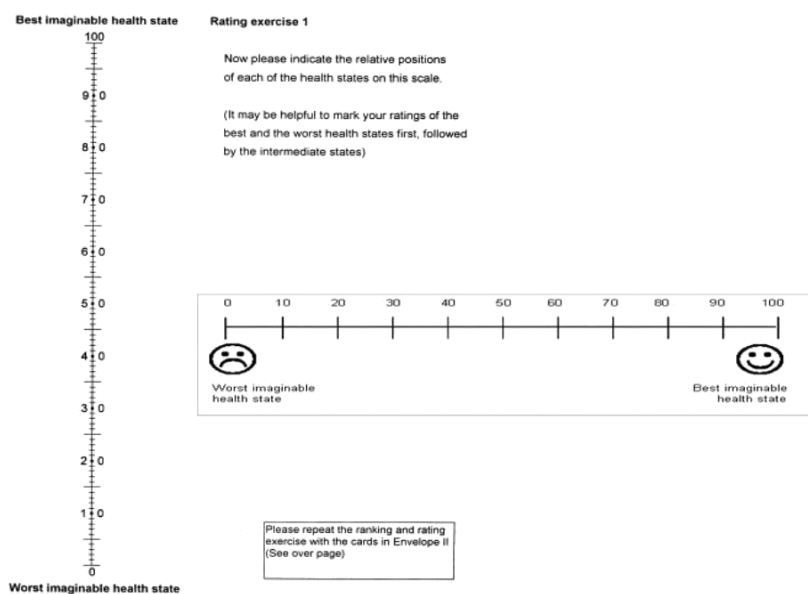
Figure II-2. Visual analog scale

Each specialization branch develops instruments appropriate to patients with specific disease conditions. As such, the score for estimating QALY by other instruments needs to be converted to values of 0 (dead) to 1 (full health). Examples are presented in **Table II-4**.

EuroQol EQ-5D to measure quality of life

The EuroQol EQ-5D uses a simple, generic approach that has been vali-dated in various countries to measure patient health status based on clinical and economic assessment. This tool was developed by a group of interna-tional researchers, including those in the pharmaceutical industry, and is officially used in HTA in the UK (NICE) and Thailand (HITAP).

The EQ-5D instrument can be formally downloaded after registration. This instrument is easy to understand and has clear instructions. The purpose of the instrument is to obtain a description about quality of life by assessing several dimensions of health status. One only needs a few minutes to com-plete the instrument or it can be conducted through a survey at a clinic, by mail, or direct interview.

Table II-4. Methods of measuring quality of life for specific diseases

Hypertension	Phisical Symptoms Distress Index (PSDI)
	The Subjective Symtom Assesment Profile
Benign Prostatic Hyperplasia	American Urogical Association Symptom Index (AUASI)
	BPH Impact index
Asthma and Allergy	Living with Asthma Questionnaire
	Life Activities Questionnaire for Adult Asthma
Diabetes Mellitus	Diabetes Specific Qol Instrument (DQOL)
Cancer	Functional Assesment of Cancer Theraphy (FACT)
	Clinically Developed Psychosocial Assesment (CDPA)
Chronic Rheumatic Disorders	Arthtritis Impact Measure Scale (AIMS)
	Toronto Functional Capacity Questionaire
AIDS	Functional Assesment of HIV Infection (FAHI)
	HIV patient-Reported Status and Experience Scale (HIV-PARSE)
	AIDS Health Assesment Questionnaire (AIDS-HAQ)

The EQ-5D instrument measures a patient's state of health. Five health dimensions are included: mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Each dimension has three levels of severity. The first level indicates no health problem present; level 2 indicates the presence of several health problems; and level 3 indicates the presence of serious problems. Thus, each health status observed has a 5-digit code, which can then be converted into a value set. This value set is a line of con-version values, in table form, on a scale of 0 (dead) to 1 (full health).

Currently, Indonesia has no referral table for quality of life, so we are temporarily using tables from Malaysia or Thailand that are accessible from websites. Thailand and Malaysia have established their own EQ-5D with a local-specific value set that was obtained from research in their country using primary data. If we compare a score of 11223 to the Malaysian referral table, the value would be 0.624, but compared to the UK referral table, that value would be 0.25. Indonesia needs its own referral table that is as representative as possible, in terms of the large variations related to societal utility spanning. The referral table is currently under development.

The instrument asks the patient to self-rate their quality of life on a "thermo-meter" of patient health status. The patient rates the value of their health status by drawing a line on a picture of a thermometer with specific numbers. This method is a subjective rating of the value of health status.

How to measure QALY

Quality-adjusted life year (QALY) is the expected final result from a health intervention, and is strongly related to the magnitude of quality of life. QALY is a combination of length of life and quality of life. The EQ-5D gives a generic measurement of quality of life.

Technically, the term "adjusted" in QALY is an adaptation of an additional year derived from utility. With this adjustment, we can obtain a number of additional years of full health. The utility value ranges from 1 (perfect health) to 0 (dead). Therefore, if a patient assesses that her/his condition after therapy is 0.8 (as result of EQ-5D) and the additional life years is 10, then the additional quality life is not 10 years, but 0.8*10 years= 8 years. The parameters of cost and utility can be adjusted by applying a discount factor within a time horizon, throughout the analysis process, in the established model (decision tree or Markov model).

Discounting

Key points
a. Cost and outcome should be adjusted over time periods, but calculated as present value
b. Specify a discount rate of 3% for both cost and outcome

Health economic evaluation results generally incorporate a long time hori-zon for interpretation. The effectiveness to cost ratio usually uses a duration of more than one year, especially for the results of therapy that are expected to last for a long term period, such as immunizations. However, society

generally expects to gain the benefit immediately, so that the parameters of both cost and effectiveness should be translated into current conditions.

In order to obtain the "present value," a correction factor is used to adjust the value. This process is known as "discounting." Several important terms related to discounting are noted below:

- Present value is to be calculated for investments over the next several years. What is the future value in terms of present value? Money promised for health care services in the future has a value lower than present value.

- The conversion processes of monetary value apply to both money paid and received in a certain time period of more than one year.

- Time value associated with money

- Discount rate for health usually ranges from 3-6%

- Discount factor is $(1+i)_t$

In order to obtain a present value, discounting or use of a correction factor is needed. This method is used to adjust future costs and benefits to the present market value.

$$PV = FVn / (1+2)_n$$

PV= present value

FV=future value

r= discount rate

n= year from start of program

The discounting rate is conceptually different from inflation level. Inflation describes price changes, however, discounting is related to monetary value considered over time. As such, value adjustment can be applied if the technology has a time range, even if the inflation rate is 0%. If the discount rate is 5% for an intervention valued at IDR 500,000 next year, the present value (adjustment for one year) would be IDR 500,000/1.05=IDR 476,190, no matter the rate of inflation.

Several experts have argued that discount rates for both cost and consequences should be similar, while others have argued that the discount rate for consequences should be lower than the discount rate for cost. Benefit is more difficult to measure than cost, so these issues are frequently debated. However, the reason for value adjustment by discounting, that the outcome must be adjusted as well as cost, is accepted by all. In Indonesian context, a discount of 3% is applied to both cost and benefit/outcome.

Table II-5. An example of discounting

Year of costs incurred	Estimated cost without discounting	Calculation	PV
Year 1	$5000	$5000/1	$5000
Year 2	$3000	$3000/1.05	$2857
Year 3	$4000	$4000/(1.05)^2	$3628
Total	**$12,000**		**$11,485**
5% discount rate			

Step 3: Building a model

Key points
a. For HTA in Indonesia, a decision tree and/or Markov model are the recommended analytic models to be applied. The preferred model depends on natural disease progression.
b. The model should be described according to clinical practices in the real world. Most importantly, a model should be developed in collaboration with clinicians, in order to gain an understanding of health/disease progression (expert opinion/ad hoc panel).
c. Parameters for modeling should be appropriate, with good data
d. Limitations of the model should be systematically reported.

HTA uses model-based economic evaluation developed according to Bayesian statistical concepts, which differ from a frequentist approach. Model-ing in the latter case is not intended to test the hypothesis. Parameters are critical to determining the appropriate modeling process. The collected data should be reliable and in accordance with the objective of the study. Essentially, modeling is beneficial for analyzing the consequences of an intervention/new technology compared to the standard, in situations that require longer time periods to observe outcomes. The relationship between phases of complicated disease progressions is difficult to follow in real life. Therefore, modeling is beneficial to "see" a result that needs longer dura-tion of time and would be too costly to conduct in real life. Hence, the model based on results from a shorter duration of time and affordable cost can be used to predict future consequences.

Modeling in economic evaluation has developed in the past twenty years. A "modeler," comprised a team of clinicians, epidemiologists, and other experts, works together to conduct health economic evaluations. Several software and spreadsheet packages are available to construct analytical and visual models, from the simplest to the most complex. Models can be con-structed with a Microsoft Excel spreadsheet, or another sophisticated software such as TreeAge®. MS Excel can be used for processing and analysis, either deterministically or iteratively, using the model's parameters (including cost, utility, effectiveness, etc.)

Decision analytical model

Decision analysis is a systematic and explicit process of applying quantitative mathematics, based on available scientific evidence. The decision analytic model applies mathematical patterns of relationship to determine the possible consequences from alternatives that are to be evaluated.

Based on model, the possibility of each consequence is expressed as a probability, and each consequence has a cost and outcome. A model is a simplified representation of a system representing the population, include-ing analytical models, visual models, or both. In health economics, modeling is conducted when it is difficult to assess a health intervention, due to time and funding limitations (e.g., estimating CEA as a lifetime

therapy) or the impossibility of conducting prospective primary research for many years onwards. The purpose of this technique is to help the analyst explore incremental cost and effect, and to assess cost-effectiveness.

Analytical model, visual model, and deterministic model

The analytical model represents a reality in mathematical or statistical form. In this model, the outcome (i.e., survival years) responds to input changes (specific risk factors) in a realistic manner.

The visual model uses symbols to represent key events, temporal and/or causal relationships, dependency, and outcome.

The deterministic model is a standard method of evaluation used with the analytic model, with one specific input set. This model can be used for decision analysis or CEA. The benefit of this model is that it provides a simpler description of a complex system that occurs in real life. The decision analytical model is expected to provide the best available evidence related to the health intervention, as well as guide the decision-making process, based on rational assumptions.

The decision analytical model provides a framework for decision-making under uncertain conditions. Several models fall under this category: the decision tree, the Markov model, patient level simulation, discrete event simulation (DES), and the dynamic model. In HTA economic evaluation, the decision tree and Markov model are frequently used; therefore, this section focuses on the application of these two models. Specifying the mathematical model in economic evaluation is very dependent on the study's objective, types of disease and disease characteristics, as well as the decision maker's objectives. Brennan et al. explained model types, taxono-my of model structures, and guidelines for choosing the appropriate model for health economic evaluation.

Probability concepts

Probability is applied to model development, in terms of expected cost and expected effect. Probability is the degree of certainty of an event in a statistical test with a value of 0 to 1. The types of probability used in model-ing are **joint probability** and **conditional probability**.

- **Joint probability** is the probability of two or more events occurring simultaneously.
- **Conditional probability** is the probability of an event if other events have occurred.
- For example, if 20 out of 30 people recovered from surgery, then the probability of recovery is 20/30 or 67%.

Decision tree

The decision tree is the simplest of decision analytical models. Alternatives or choices from each intervention are explained in branch form, as shown in **Figure II-3.** The first node of the branch in a decision tree is a square. This node is a decision node, explaining the decision question. For example, drug A or drug B, or, perform screening or not? The branch after the decision node gives two option nodes (drug A or drug B, or, with or without screening). The figure shows only two interventions, but it is possible to add more option nodes for other types of intervention.

In **Figure 11-3**, the next branches are *chance nodes*, represented as circles. These nodes are the possible events for each alternative intervention given, after the decision node. Every alternative should be mutually exclusive (only one option). The probability value from all events should be = 1 (one) or 100%. For instance, if the probability of a cancer event is 0.6, then the probability (p) of cancer absence would be 1-p, or 1-0.6=0.4, or 40%.
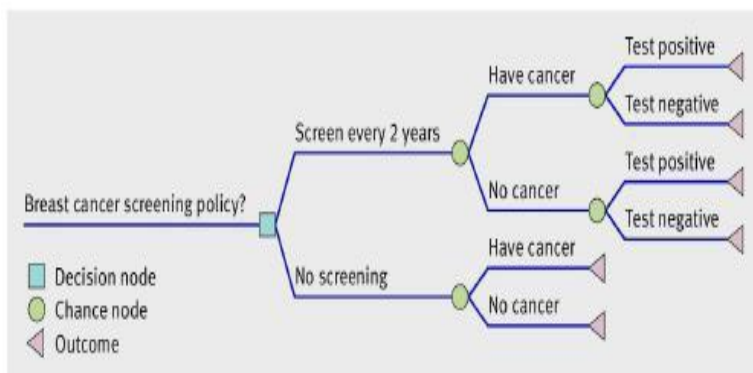
Figure II-3. Decision Tree
Source: Petrou and Gray, Economic evaluation using decision analytical modelling: desing, conduct, analysis, and reporting, 2011.

The last nodes of the branches are the terminal nodes, represented as triangles. If there are no further branches (or events) after these nodes, then they are considered to be the end-points. The benefit of a decision tree is its simplicity and transparent structure, as alternatives can be clearly seen at the beginning of the model. However, this model is limited by the variable of time, and difficult to apply to long-term interventions (i.e., time depen-dent). Recurrant events cannot be easily included in a decision tree. For chronic diseases, a decision tree would have long "bushy" branches, lead to confusion, and probably not adequately answer questions related to the probability of chronic disease transition and progression.
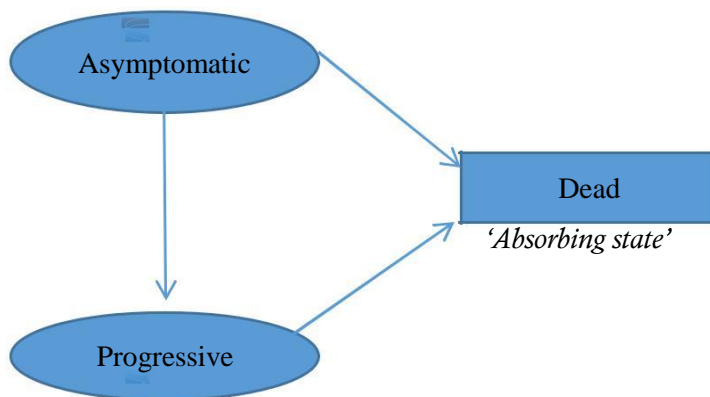
Markov model

Although a decision tree can be used to explain the sequence of disease transition events in a certain time of period, it cannot be used to differen-tiate between prior and future events (changes in disease severity). In addition, the decision tree is not flexible enough to explain recurrences. Hence, a decision tree is inappropriate for chronic illness models, in which a sequence might include complications, recurrences, recovery, or even the probabililty of death. Complex conditions are difficult to accurately portray in a decision tree, since complexity by nature necessitates the use of many branches. Hence, the Markov model is would be more appropriate for chronic illness.

Markov modelling includes variations such as the Markov model, Markov chain, and Markov process. The Markov model is a stochastic process involving continuous random changes in which the next future event is independent from past events (e.g., the fluctuative share market, currency changes, or a coin toss, which cannot be predicted).

The Markov process is related to finite health states, in which the transition or change from one state to the next has a probability value. It is assumed that the state of a sick patient depends on his present state. As such, the model can be used to analyze worsening, improvement, or recurrence. The *Markov chain* is a unit of the Markov process describing a patient's state in discrete time, that can change due to sickness transitions or patient health.

On the Markov diagram (**Figure II-4**), a bubble depicts a state of health. The arrows denote the transition between states. Since it is impossible for a dead patient to transition back to a healthy condition, the "dead" state is called an *"absorbing state."*

Figure II-4. Markov model description

Markov models include the parameter of the probability of changes in health state (i.e., stage of disease is a description of the clinical outcome), cost, and effect. Markov models have the following characteristics:

- The model describes transitions between patient states with particular time intervals and clear time dimensions. For example, Asymptomatic → Progressive → Dead

- Model development is done according to specified time sequences between stages, also known as cycles.

- The speed of disease severity is assessed based on probability value from one stage to the next in the disease progression (tran-sition probabilities). The probability of asymptomatic to progres-sive to death can be calculated and analyzed to estimate the changes in cost and effect.

- Cost and effect from each state and/or transition condition is calculated and analyzed.

Phases of model construction:

- Sample data of different health states are analyzed in the model and should be mutually exclusive, representing important clinical conditions and costs in the disease transition process. The ISPOR Task Force stated that the model must reflect our biological and theoretical understanding of disease for CEA.

- Determine the observation cycle for each intervention.

- Determine the total cycle.

- Identify the number of disease transitions that receive intervention.

- Identify probabilities of each disease transition phase with its intervention (medicine), (i.e., transition probabilities).

- Identify the cost and effect of each health state and transition (the change from one state to another)

The duration of model cycle should be clinically justifiable, and short enough to represent the frequency of clinical events and interventions. The shorter the cycle, the higher the precision, but efficiency (CEA) result will decrease. All events should be included in the analysis. The time horizon to assess the intervention on the progression to another health state should be long enough to observe an effect on disease transition and the cost. For example, the patient lived for 70 years, which was his life expectancy.

The ISPOR Task Force recommends that transition probabilities and effects of interventions should be obtained from good quality data and sources. The analysis must synthesize data from clinical test results and other sources, according to the evidence search model: 1) the relationship be-tween clinical decision or intervention and health effects, 2) the health provider responsible for the decision or intervention, and 3) the target population. If the evidence search result reveals a disagreement in parameter estimation, then it can be reconciled through a meta-analysis.

Modeling example: In a hypothetical case, the parameters are as follows: the probabilities of asymptomatic patient status to progressive = 0.1667, and

progressive to death = 0.3333. The cost of treatment for the progressive state using a new intervention or medicine is Rp 3 million and its utility (QALY) is 0.76. Using *Excel* software or *TreeAge*, the analysis is done, followed by a sensitivity analysis. The obtained result is an ICER value, which can be compared to a threshold in order to decide between the old and new interventions (**Table II-6** and **Figure II-5**).

Table II-6. Example of Markov model parameters

Probability	Asymptomatic	Progressive	Dead
Asymptomatic	0.6667	0.1667	0.1667
Progressive		0.1667	0.3333
Dead			1.0000
	Cost of the progressive state	Utility of the progressive state	
New intervention	Rp 3 million	0.76	
Old intervention	Rp 1 million	0.65	

Utility of asymptomatic state

Cost of asymptomatic state
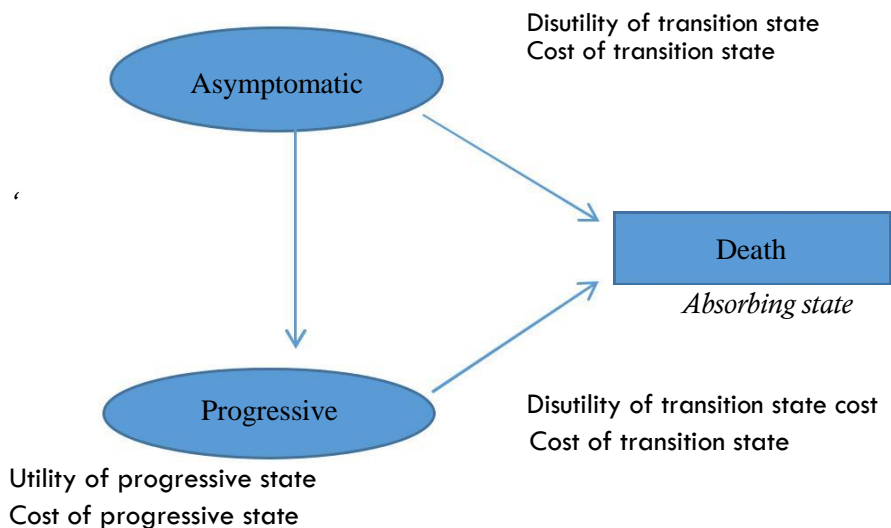
Figure II-5. Markov model with parameters

Step 4: Sensitivity analysis

Key points
a. A sensitivity analysis is used to explain uncertainties in the measurement paramaters.   b. Choose the appropriate method of sensitivity analysis: one-way, multi-way, or probabilistic.

Sensitivity analysis (SA) is a way to analyze uncertainty from an economical analysis or decision analysis. SA is important because model paramaters such as cost estimation, effectiveness, and cost effectiveness from economic evaluation results have some level of uncertainty, due to several reasons, mainly from uncertainty of the input parameters. For example, data analy-sis that uses a mean has limitations if the data are not distributed normally. In addition, extrapolation from one point to another, possibly over a long period of time, may lead to uncertainty since the actual observation was done over a short time period. Combining and generalizing study results also contributes to uncertainty. Therefore, this uncertainty should be evalu-ated and reduced as much as possible. The most influential parameter and its effects on study results should be identified, quantified, and interpreted.

Generally, there are two types of sensitivity analyses: deterministic sensi-tivity analysis (DSA) and probabilistic sensitivity analysis (PSA). DSA is conducted using a specified input model consisting of estimation points/values that vary according to plausible assumptions. PSA is the analysis of all parameters using a distribution of various values, in a simulation approach. Three types of DSA are the *one-way, two-way*, or *multi-way sensi-tivity analysis.*

The simplest and the most frequently-used DSA in health outcome studies is the *"one-way"* method, which changes the value of one variable using a plausible value range, while other variables are kept constant. The variables are quantified and analyzed. For example, does the *discount rate* for cost and effect in economic evaluation affect the analysis result for long-term time

horizons? It is possible to choose a baseline of 0% (no discount), 2%, 3%, 4%, 5%, and 6%. So, in a *one-way* SA, the discount rate would be 0%-6%, then the study result (running model result) observed is the incremental cost per life year saved in treatment A, compared to treatment B. One-way SA can also be used to explore the robustness of a study result using a developed model (model power).

The Tornado diagram is a SA in the form of a graph. Generally, the long bar in the chart appears at the top and other variables below are ordered by ranking or the level of influence on the effect (sometimes affecting the ICER). The graph is shaped like a tornado. The disadvantage of *one-way SA* is simplification, since real states are complex, with many variables which probably interact or influence one another.

In a Markov model, it is assumed that a patient remains in a health state for a cycle, such as for 1 year, followed by a transition occurring at the end of the cycle. In fact, the patient's state of health is changeable at any time in the cycle. So, the assumption that a patient will remain in a certain state for a year is not representative of reality. Consequently, cost estimation and effect becomes overestimated. Analysts attempt to balance the overestima-tion using a *"half-cycle correction,"* for example, mathematically dividing the cost result value and effect of the first and the last cycle into half.

Researchers frequently face the question: does a study which collected primary data with a good epidemiological design need modelling? The answer is **yes,** if the study was done for a short time and the result will be extrapolated according to a time horizon to achieve the goal of proving the long-term effect of the intervention. For example, a study is done to assess the effect of an anti-hypertensive medicine on heart failure. The measured outcome is an intermediate outcome, controlled blood pressure. Because the goal of CEA or CUA is to explore the long-term effect on patient health status, the study result alone is not enough to convince the decision-maker, as he expects the assessment to continue until effectiveness (final outcome) can be measured. Modeling will provide an extrapolation of the interme-diate results, allowing the researcher to more adequately answer the research question. Studies using primary data to prove clinical effectiveness are usually expensive and require long periods of time.

Efforts to prove that our model is robust through SA is important for both patient data analysis and model analysis. The issues to be considered for SA are as follows:

1. Identify uncertain parameters

All variables are potential candidates for SA. A parameter can be excluded from SA if it indeed has *"absolute certainty,"* or if the initial analysis result shows that even when the parameter varies, it has minimal effects on study results.

2. Specify a plausible range

Acceptable range values are obtained from:

- Careful literature review

- Consulting the opinions of relevant experts

- Using confidence intervals (CI) of the mean (for stochastic data)

We should be careful in assessing experts' justifications regarding the range probabilities. The scientific journal authors sometimes state exaggerated opinions related to the analysis, which may lead to doubt. For example, an author may state that the result was very strong (robust), although the estimation range of the key variables was very low. In addition to using point estimates for base-case (best guess) estimates and upper-lower bounds, one can use probability distribution for range specification.

### 3. Decide on the method (DSA or PSA) and perform the sensitivity analysis

Examples of sensitivity analyses.

The *one-way sensitivity analysis* reflects variations in parameter values that influence the model result. This technique reveals a change in only one paramater at a time. This technique can be used to assess a parameter perceived to be the key driver of model result.

For example, a modeler examines which parameter has the most influence on a model result, then he assumes a specific change in the key parameter

value. For instance, all parameters increase or decrease as much as 20% from the original value. For every parameter, the modeler records all percentages that influence the evaluated outcome. SA results can be seen in a Tornado diagram.

Figure II-6 shows the changes occurring in response to an increase or decrease of the parameter value by as much as 20%. The significance of the change should be noted, with regards to a change in ICER (increase). The gap between the before and after results indicates the level of uncertainty of the model. For instance, if the probability value for success of an intervention decreased as much as 20%, then the change in cost-effect-iveness ratio could be more than 10%. Or, if the discount rate decreased as much as 20%, then the change in cost-effectiveness ratio could decrease as much as approximately 20% from the previous result.
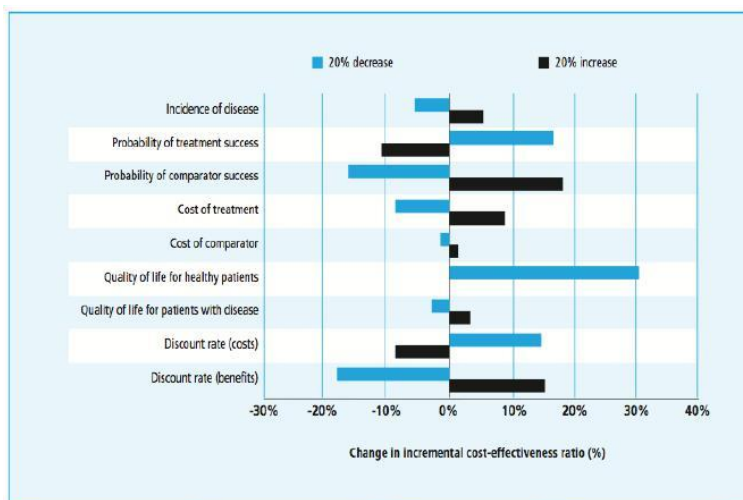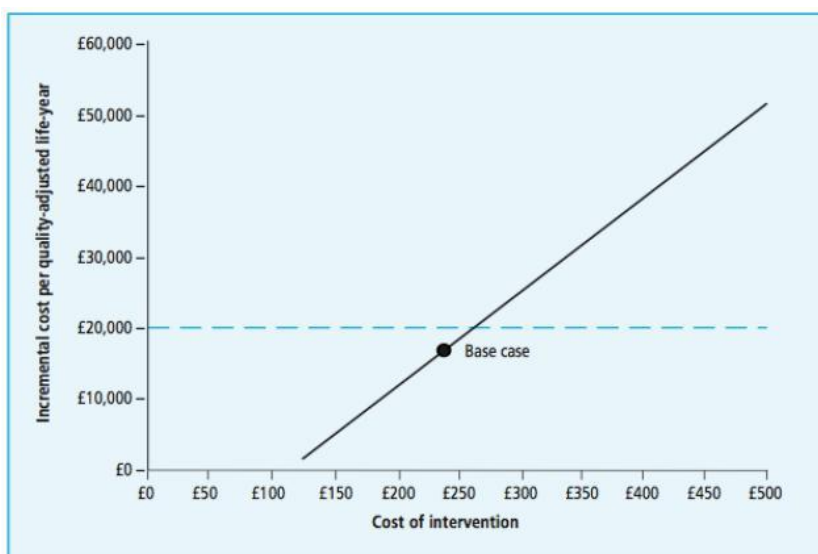

Figure II-6. Tornado diagram
Source: Taylor, What is sensitivity analysis? Available at:
www.whatisseries.co.uk, Hayward Medical Communication, 2009.

Although a Tornado diagram is useful for showing the effect of a parameter on the outcome, it is less helpful for illustrating the confidence level of the model. If the parameter's confidence level is very low, it is reasonable to think that one of input parameters is less representative (e.g., unavailable data). For example, data about a drug's effect on patients' long-term

mortality is needed. If the available drug data is only for one month, but the data of another parameter is for 20 years, the standard error value will be affected.

Furthermore, this technique can illustrate the highest and lowest possible values of a parameter. The definition of 'possible' in this context could be different from the other models, but at least the confidence interval value and number range could be presented in the model by comparing them to other studies or publications.

The next analysis to be discussed is the threshold analysis. This analysis is used to compare the value derived from an economic evaluation to an acceptable limit value (the ability of a country to bear the cost or willingness to pay/WTP). For example, the threshold value in the UK is set at £20,000. An example of a threshold analysis result is shown in **Figure II-7**.

Figure II-7. Example of a Threshold Analysis
Source: Taylor, What is sensitivity analysis? Available at:
www.whatisseries.co.uk, Hayward Medical Communication, 2009.

Figure II-7 illustrates that the ICER would be below the threshold (£20,000) only if the intervention cost is below £270. If the cost of the intervention increases, cost-effectiveness will decrease or go above the threshold.

The *multi-way sensitivity analysis* explains the relationship between two or more parameters. For example, the two key parameters of cost and outcome produce a combination value in a given range. Interpreting the multi-way sensitivity analysis can be complicated because of the many parameters in modeling. A commonly used method to evaluate the confidence interval of a parameter is by taking a close look at extreme values, i.e., studying the variations of all parameters by best- and worst-case scenarios, chosen based on the perspective of an evaluated intervene-tion. For example, in a scenario with the most optimistic value, a parameter is chosen based on the CI value related to its input.

Figure II-8 shows a two-way sensitivity analysis of the parameters of effectiveness and price of intervention. If effectiveness increased to 40%, the price would reach £700. This analysis shows that ICER values fluctuate, until one is clearly below threshold <£20,000 and one is above the maximum limit. The results show that if effectiveness increases to 40% and the price is £500, the ICER will be £13,165.

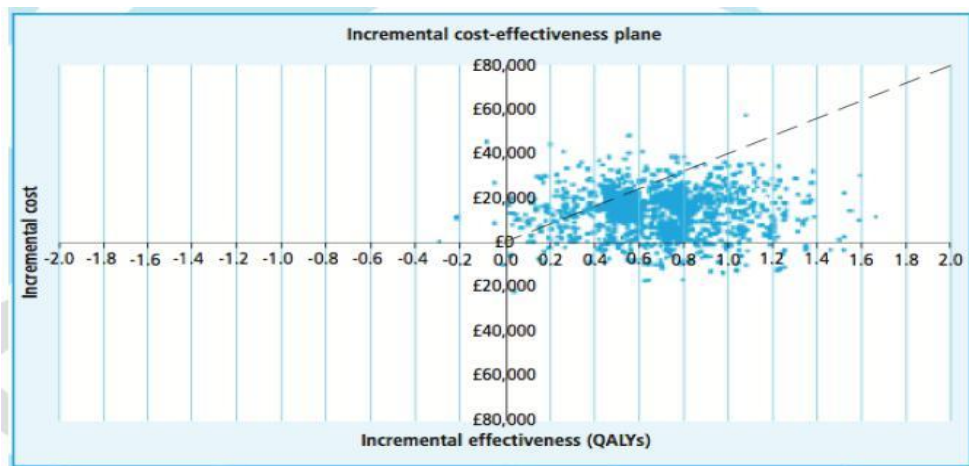| Table 1. Two-way sensitivity analysis								
	Price of intervention							
	£0	£100	£200	£300	£400	£500	£600	£700
0%	Cost saving	£4,662	£7,459	£11,935	£19,096	£30,553	£48,885	£78,215
5%	Cost saving	£4,200	£6,720	£10,752	£17,203	£27,525	£44,040	£70,464
10%	Cost saving	£3,780	£6,048	£9,677	£15,483	£24,773	£39,636	£63,418
15%	Cost saving	£3,402	£5,443	£8,709	£13,935	£22,295	£35,673	£57,076
20%	Cost saving	£3,062	£4,899	£7,838	£12,541	£20,066	£32,105	£51,368
25%	Cost saving	£2,756	£4,409	£7,054	£11,287	£18,059	£28,895	£46,232
30%	Cost saving	£2,480	£3,968	£6,349	£10,158	£16,253	£26,005	£41,608
35%	Cost saving	£2,232	£3,571	£5,714	£9,142	£14,628	£23,405	£37,448
40%	Cost saving	£2,009	£3,214	£5,143	£8,228	£13,165	£21,064	£33,703

Effectiveness of intervention

Figure II-8. Example of Two-Way Sensitivity Analysis
Source: Taylor, What is sensitivity analysis? Available at:
www.whatisseries.co.uk, Hayward Medical
Communication, 2009.

Probabilistic sensitivity analysis (PSA). In most models, every parameter is presented as point estimate value. For example, the probability of success of an intervention is 60% with 95% CI between 42% and 77%. In PSA, in addition to choosing a value for every parameter, the distribution of parameters in the model is also analyzed, with the assistance of a computer application (simulation or running model). The range is set by mean value, standard deviation, and the form of data distribution. The parameter values should always be noted, the probability number should always be between 0 and 1, and the cost should never be negative. Every time the model simulation runs, the computer application will randomize the parameters, record the results, and illustrate the variations (usually a randomization/ running model is done 1,000 times or more). The analysis result of this software application is presented in the form of a cost-effectiveness plane, with the value distribution in the form of a scatter plot. Every dot shows the incremental cost as well as the incremental effectiveness of the inter-vention.

Figure II-9 shows that input has a relatively high uncertainty. The model result shows a quite wide incremental distribution. In contrast to the next figure of the same model, if the confidence level of an input is higher, the distribution will be denser and narrower. We should keep in mind that the mean of every model is identical but with different confidence levels. As an illustration, a threshold line in SA could be elevated. In the example, using the UK's perspective value, the threshold is < GDP £30,000. **Figure 11-9** shows high uncertainty, because of the wide distribution of the scatter plot. Some dot values are above threshold (broken line), while others have decreasing QALY (negative numbers) and cost distribution that reaches £40,000, even though there are some negative costs. Although we see that the QALY increase is in accordance with costs, the model has a high uncertainty. Therefore, interpretation and information usage should be done carefully.

An alternative analysis to obtain a better description of results, is the cost-effectiveness (CE) plane. **Figure II-10** shows slight invisible uncertainty, a dense plot under the threshold line, and that the intervention has great potential for being cost effective. As QALY increases, the cost incremental

value remains around £20,000, not increasing much. However, some plot points have values above the dashed line. After reading the cost-effectiveness plane result, we must analyze a cost-effectiveness acceptability curve (CEAC) (**Figure II-10**). The x-axis in CEAC shows the range of willingness to pay (WTP) values, from lowest to the highest. The y-axis shows the opportunity of the intervention to be cost effective.
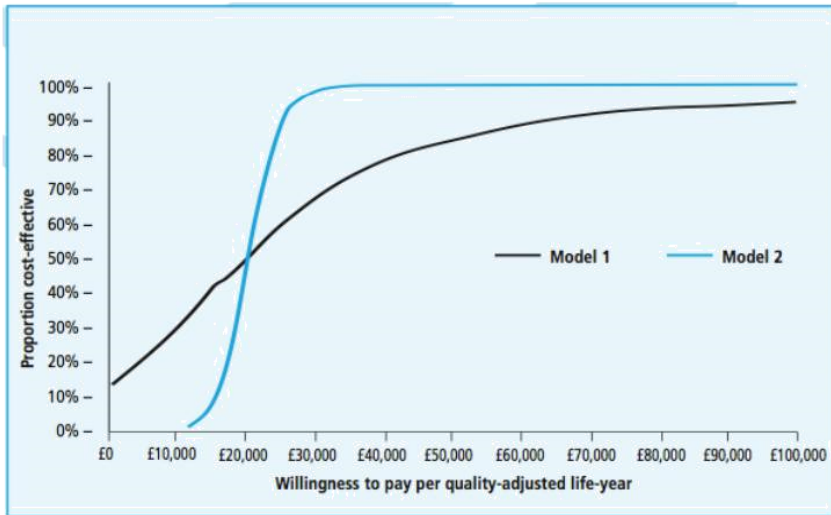

Figure II-9. Example of a Cost-Effectiveness Plane (A)

Source: Taylor, What is sensitivity analysis? Available at:
www.whatisseries.co.uk, Hayward Medical Communication, 2009.

Model 1 of the CEAC shows that for WTP of £30,000/QALY gained, the opportunity to be cost effective is 67%. However, model 2 has an opportunity of 98%. Although the mean values of the cost-effectiveness plane in both models are the same, the decision-maker can have more confidence in the second model, because it has much higher opportunity or probability to be cost effective.

To interpret the PTK study well, the results of the CEA or CUA study must be interpreted correctly, including how ICER is obtained, interpreted, and compared to a threshold. The model analysis results are in the forms of a cost-effectiveness plane and a cost-effectiveness curve, which are both used to describe the results of our economic evaluation.

Figure II-10. Illustration of a cost-effectiveness acceptability curve (CEAC)

Source: Taylor, What is sensitivity analysis? Available at: www.whatisseries.co.uk,. Hayward Medical Communication, 2009.

Step 5: Result interpretation

Key points
a. The result interpretation describes if the research purposes were achieved and if the proposed health technology was proven to be cost effective.
b. The comparison of the ICER value to the threshold allows the decision maker to determine if the proposed intervention has "value for money."
c. The SA result is necessary for describing parameter uncertainty in the model.
d. The study results must be interpreted carefully with regards to the study limitations.

Each country has a threshold to decide whether the proposed intervention results (ICER) is cost-effective, based on the willingness to pay (WTP). For example, NICE in the UK has a threshold of £20,000 – £30,000. The WHO

suggests that developing countries without a threshold value to use GDP per capita (1x GDP per capita is considered to be "very cost-effective" and 3x GDP per capita to be "cost effective"). Thailand set its threshold at "1.2 x GDP per capita" or around 120,000 Thai Baht. In the context of Indonesia, since no threshold has been determined, we will use the GDP per capita criteria that has been adjusted for purchasing power parity (PPP), until we have our own threshold value / WTP.

Results from a threshold analysis to assess cost effectiveness of a proposed intervention, can be narrated as follows: "The ICER of…(option A)…is less than...(GDP/capita) as long as…(parameter X)…is more than...". Or it can also be: "…(option A) is more cost-effective than...(option B) as long as...(parameter X) is higher than… ".

Example of result interpretation of an economic evaluation

The following is an example of result interpretation of economic evaluation on sildenafil for pulmonary arterial hypertension therapy in Indonesia.

- Sildenafil significantly improves functional class over placebo. In contrast, no significant differences in functional classes II and III were found between patients given beraprost vs. placebo. More studies are necessary in order to obtain accurate conclusions.

- The economic model estimates that sildenafil will increase the patient's life by 1-3 years (life years gained) over beraprost.

- **Table II-7** describes the cost, life-years gained, and QALY for the two types of therapy in functional classes II and III. **Figures II-12 and II-13** describe the SA results: the CEAC shows that sildenafil has greater opportunity to achieve cost-effectiveness (potentially cost effective) than beraprost.

- The economic evaluation model result shows sildenafil to be the recommended therapy option for PAH patients in Indonesia, compared to beraprost, the current therapy available in Fornas). The ICER value compared to the Indonesian threshold using GDP per capita (Rp 43 million) results in an ICER per QALY under the threshold. Hence, sildenafil has good value for money.

Generalization of study results for the purpose of testing the hypothesis depend on design and sample collection technique. For example, is the sample from hospitals in Java representative of Indonesian hospitals? Keep in mind that the model is not hypothesis testing, and parameters analyzed are not intended to determine a causal relationship. For those reasons, economic evaluation in HTA can be done even with limited sample size. For each parameter measured, the minimum sample size must be fulfilled; however, the relatively large number of subjects needed in hypothesis testing studies is not required.

Study limitations should be discussed, such as samples that are not representative of the Indonesian population, methods used, input parameters, etc. A description of components and cost variation will be noted by the stake-holders who wish to improve cost standards for fairer resource allocation in the various regions. Economic evaluation data of a new intervention using a modeling approach can be used to generalize representative cost and output data. As such, the model can provide information as to the potential of an intervention having value for money for the country, if it is selected as a policy option.

Table II-7. CUA study results of sildenafil compared to beraprost

	FC II		FC III	
	Bera	Silde	Bera	Silde
Total lifetime cost (IDR x10$_3$)	496,000	520,473	426,077	484,879
Life-years gained	16.23	16.94	14.1	15.8
QALY	11.9	12.47	10.08	11.58
	FC II		FC III	
	Silde vs. Bera		Silde vs. Bera	
Additional cost (IDR x10$_3$)	24,472		58,801	
Additional QALY	0.57		1.51	
ICER per QALY (IDR x10$_3$)	42,843		39.02	

Source: Study report of PTK Ministry of Health, 2016. Note: FC: functional class
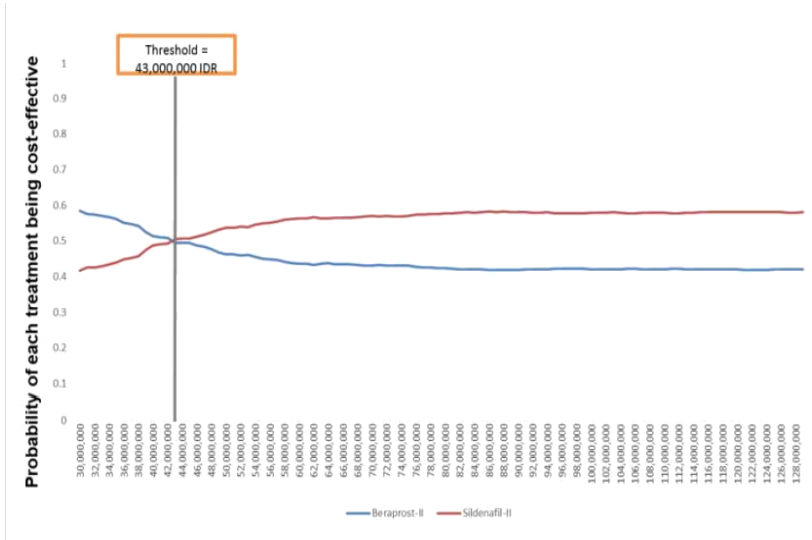
Figure II-11. CUA result of sildenafil SA for PAH therapy compared to beraprost for functional class II
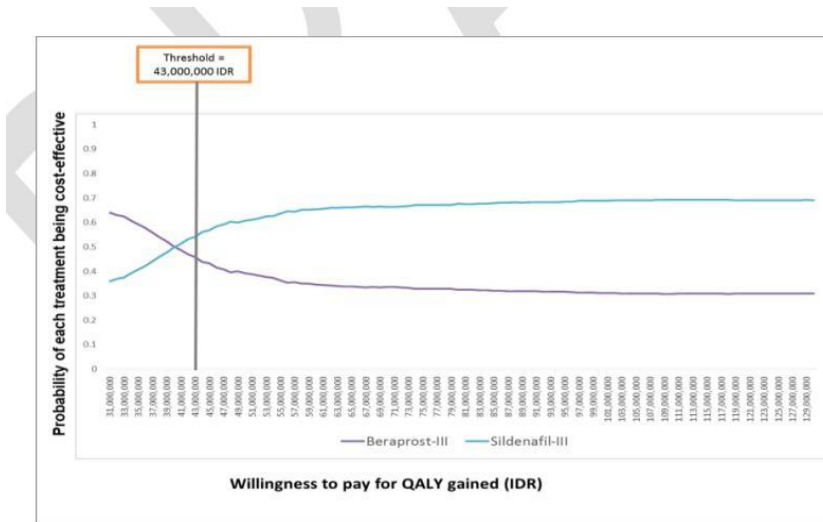
Figure II-12. CUA result of sildenafil SA for PAH therapy compared to beraprost for functional class III Source: Study report of PTK, Ministry of Health, 2016.

126

Step 6: Reporting

Key points
a. Clear and accurate reporting of PTK study results is a reflection of the quality and transparency of the study b. All components of the study must be reported, but important information should remain the focus.

It is important to use the proper writing format on the HTA economic evaluation study result report for several reasons. First, using proper report format reflects on the study's transparency. Clear and accurate reporting allows readers to easily understand data collection techniques, how the analysis was conducted, study limitations, and more. Second, a well-written report can be easily compared to other study results. As such, the readers will find no reason to doubt the study results concerning an intervention, especially in cases of new drugs proven to be cost-effective according to the study results from a correctly-designed analysis, compared to other study results. Publishing the study results is also recommended.

Important items of the evaluation should be noted in the report of the study results, similar to a reviewer conducting an appraisal of a study. While clinical studies only provide information on the output, an economic evaluation also covers other components such as resources, cost, utility, and cost-effectiveness. Economic evaluation studies may prove challenging to edit-ors, reviewers, and anyone using the study results. As such, clear stan-dards are necessary to assess the quality of the results. The **Consolidated Health Economic Evaluation Reporting Standards (CHEERS)** is a refer-ence for economic evaluation study result reporting. It was developed by the ISPOR expert panels. Seven main items are to be considered, namely:

1 Title

2 Abstract

3 Background and purposes

4 Methods, including

a. Population and sub-group targets

b. Study location and place

c. Study perspective

d. Comparison

e. Time horizon

f. Discount rate

g. Measured health output

h. Effectiveness and its measurement results

i. Measurement of preference output

j. Cost estimation (summary on data collection and cost description

k. Exchange value, price and conversion (the adjustment to current values / exchange value, etc.)

l. Model selection / modeling

- Assumptions used

- Analysis method (abnormal distribution, missing values, extrapolation, etc.)

5 Results, including modeling results, (parameter, ICER, characteristics of study uncertainty, heterogeneity / variations in cost, output, or cost effectiveness).

6 Discussion, on the study results, limitations, study generalization, and relatedness to current conditions

7 Others: sources of funds, conflict of interest statement

The **check-list according to Drummond,** includes statements concerning descriptions of:

1 Researchers' questions

2 Measured intervention alternative

3 Measurement on intervention effectiveness

4   Measurement on cost and consequences

5   Results of cost measurement and consequences, which are accurate and based on the determined units

6   Credibility of the measurements of cost and consequences

7   Cost and consequence adjustments over time

8   Incremental analysis for cost and consequences

9   Handling uncertainty in cost estimation and consequences

10 Delivery and discussion of study results including all the important issues, such as ratio of cost effectiveness, differences from other study results, generalizations, possibility for implementtation in uncertain conditions such as fund limitations, etc.

Chapter 12
Budget Impact Analysis

The importance of budget impact analysis

Key points
a. HTA requires budget impact analysis (BIA) after the economic evaluation is performed.
b. BIA should also be determined even if the health technology is found to be not cost-effective, but included in the JKN benefit package, because it is needed by the community.

Objectives of budget impact analysis

After a proposed health technology has been investigated and determined to be cost-effective for both output and economic aspects, a budget impact analysis (BIA) should be performed. Since the objective of the BIA is to explain the implications of adopting the new health technology, it has to be performed from the payers' perspective, in this case, BPJS for Health or the Ministry of Health (MoH).

BIA is an important component of a comprehensive economic evaluation on health technology. The BIA and the cost-effectiveness analysis (CEA) are required before the benefit package can be determined. Technically, the BIA explains the probability of both treatment and cost alteration of certain diseases.

BIA estimates financial consequences from adopting and diffusing new technology (medicines or medical equipment) over a certain period of time. According to ISPOR (International Society for Pharmacoeconomics and

Outcomes Research) guidelines, *BIA is defined as the best possible estimation of the financial consequences for the budget holder resulting from the adoption and diffusion of a new pharmaceutical drug or medical device over a well-defined time period.*

BIA is also needed to devise a budget, assess future needs (forecasting), and make decisions. As such, it needs to be tangible, rational, and reliable in identifying significant factors in the analysis, including any uncertainty.

How does an innovative treatment affect future budget plans?

- A new treatment might have a higher cost than the current ones

- A new treatment might significantly reduce the cost of illness, because it is clinically beneficial for patients

- A reduction in cost of illness might balance a higher cost of future treatment

- Timing of alteration intervention and cost of illness often affect the budget.

Why bother estimating the BIA output?

- To better estimate treatment cost in order to achieve certain goals that affect the budget

- To inform decision-makers about future intervention-related health service planning, such as fluctuations in visits to hospitals/clinics, length of stay, etc.

- To inform decision-makers about a health benefit when BIA is implemented, to justify the proposed budget

Correlation between BIA and economic evaluation

BIA is not an economic evaluation, however, it is an integral part of economic evaluation results of a proposed health technology or intervention. If the result showed that the proposed health technology was considered to be cost-effective at a current or proposed rate for the benefit package, then BIA should be conducted. However, if the result was nega-

tive (found to not be cost-effective), then a threshold analysis needs to be done in order to assess if the lower price has value for money. This particular price is potentially useful for negotiation and is an important reference for BIA.

If there were not enough cases in the technology assessment (e.g., treatment for rare diseases), then the proposed technology was considered to not be cost-effective. However, it is still critical to develop a BIA to provide evidence for decision-makers who need to improve the benefit package for vulnerable groups of patients with rare diseases in Indonesia.

BIA Method

Key points
a. BIA has a different objective from economic evaluation, though BIA can be done with a similar template.
b. BIA needs additional data to estimate coverage, including prevalence and incidence rates, the population number, and cost, using a payer perspective.
c. BIA is commonly done within five years without discounting.
d. The BIA result is presented with the economic evaluation results

Differences between BIA and CEA

BIA is different from CEA; BIA is not used to assess whether a proposed health technology will be beneficial for people and provide value for money (the result is worth the effort). **Table II-8** below shows several aspects that distinguish BIA from CEA.

Table II-8. Differences between BIA and CEA

Methods	Population	Time Frame	Sample of Assessed Output	Value for Decision Makers
CEA	Annual or individual cohort data of incidence rate for particular case	Referred to certain patho-physiology period of diseases	Incremental life years, Incremental QALY cost/QALY gained	Inform decisions on dominant intervention as a result of comparing to other options (provide value for money compared to the threshold)
BIA	The whole population/ estimated suffering population	Annual base for five years ahead	Annual trend of service cost's alteration within five-years Annual trend of mortality and morbidity rate within the same time frame of BIA	Budget planning Output achievement (program/ intervention)

Required Data

The BIA includes a calculation of the observed intervention cost (e.g., drugs), estimation of service fee deviation compared to the existing intervention, as well as the impact on budget escalation and sensitivity analysis (SA). The components analyzed determine what kind of data is needed, namely:

- Output: all estimated expenditures and cost savings are assumed to be associated with the effects of comprehensive health services around the country (from a national perspective). A specific per-

spective might be related to the drug alone, or the drug procurement cost impact, but ideally (recommended), the perspective would cover the entire impact on health care costs, as a result of including new interventions. Adjustment by discounting is not needed because BIA is intended to explain the implications of funding / budgeting for additional needs, as a consequence of developing a new benefits package.

- Health status and target population: The health condition of the population as well as existing treatment patterns should be comprehensively explained in detail and in relation to interventions for observed health problems. Estimating the potential target population with access to interventions (e.g., medicine) includes the entire population (patients) eligible to obtain a new drug in a certain time period. Therefore, prevalence rate is required in addition to incidence rate, because patients who previously had access to available treatment (interventions) would likely seek and obtain the new drug/ intervention. This scenario is known as induced demand, and it promotes market expansion of the new drug. Increased demand can also be due to increased number of patients.

- Introduced intervention and standards of comparison: The efficacy, safety, effectiveness, and side effects of new interventions/ health technologies/drugs need to be compared to those of current interventions/drugs. These aspects of the comparison should be fully elaborated, as their impact will be measured against the proposed intervention. The impact is reflective of the analyzed factors, including incidence rate of diseases treated with observed drugs/interventions, diagnosis, treatments, other resources, and cost.

- Time horizon: The duration of assessment must be approved by the decision-makers. It is advisable to use a five-year observation as the base case, and it is obligatory to present the flow of funds needed for each year.
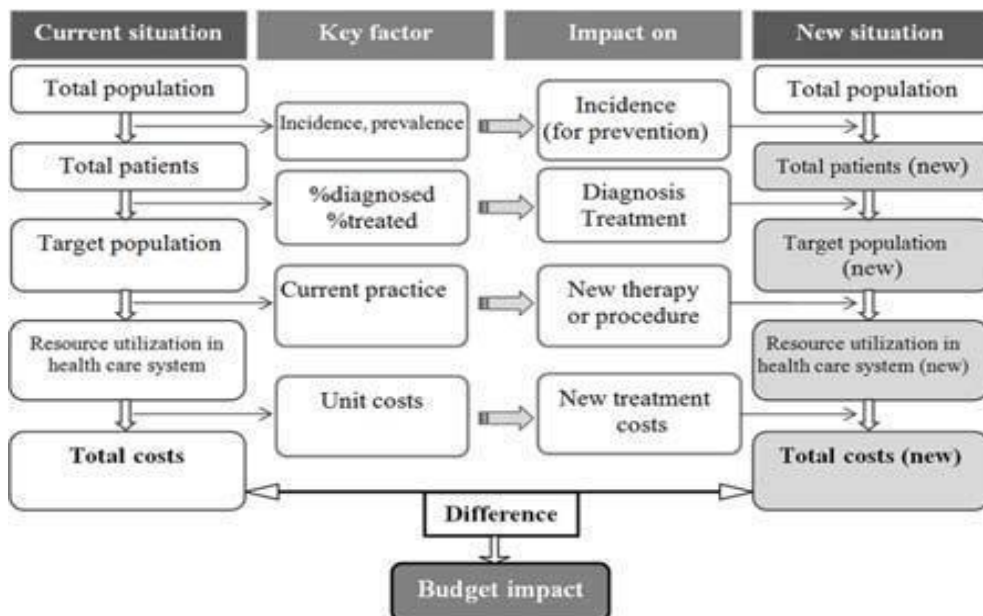
Framework of BIA

Technically, BIA consists of six (6) stages as follows:

1 Define a population as well as its characteristics

2 Select a time period/horizon

3 Compare current treatment/intervention with the following year's treatment using new interventions/technology

4 Estimate the intervention/health technology cost (e.g., drugs)

5 Estimate the changes in economic burden of disease

11 Present the results

The framework of BIA was first introduced in 1998. Since that time, many countries have performed BIA as part of HTA and used the results to support decision-makers in determining the list of drugs essential to the benefit package.

Examples of BIA frameworks implemented in Thailand and Taiwan:

Figure II-13. Framework of BIA implemented in Thailand
Source: Journal Medical Association of Thailand (2014)

BIA working tool					
Current market size (in past 2 years)	Number			Data source	
The size of patient population (n/year)					
The size of patients taking treatments (n/year)					
cost of the treatments ($)					
Expected future market size (in oncoming 5 years)	1st	2nd	3rd	4th	5th
Total drug cost of the treatments ($)					
New treatment				Data source	
(a) Monthly drug cost (per patient)	$/month				
(b) Treatment duration	months				
Year	1st	2nd	3rd	4th	5th
(c) Predicted annual number of new patients who will take new drug					
Total person-time on treatment (including new and current users)					
(I) Annual drug cost					
Substituted existing treatment				Data source	
(d) Monthly drug cost (per patient)	$/month				
(e) Treatment duration	months				
Predicted annual number of new patients who suppose to take existing drugs	same as (c)				
(II) Annual saving drug cost					
Financial impact = (I) – (II)					
For sensitivity analysis					
scenario	1st	2nd	3rd	4th	5th

Figure II-14. Framework of BIA implemented in Taiwan
Source: ISPOR 2012

Sample BIA of a study result in Indonesia

The application of BIA leads to policy recommendations on the inclusion of a health technology into the benefit package, which are adjusted within the

context of the country, taking into account the prevalence rate, coverage, and other important information.

Sample 1: BIA on the Policy of PD First (Continuous ambulatory peritoneal dialysis as the first option to treat renal failure)

The economic evaluation result of hemodialysis (HD) compared to continuous ambulatory peritoneal dialysis (CAPD), commonly referred to as peritoneal dialysis (PD), recommended wider coverage of PD. The policy of using PD first meant choosing PD as the first medical intervention for patients without complications or contraindications. This PD First Policy was conducted to improve patient equity for treating kidney failure, in order to increase coverage, as it was challenging to provide HD widely, particularly in remote areas. The BPJS and Ministry of Health were expected to provide sustainable infrastructure and fair reimbursement to the providers. However, HD could be used as the first option for patients with particular medical conditions.

To perform a BIA on the PD First Policy, we used the provider's perspective to calculate the amount of resources (including cost) needed to conduct treatments within a 5-year time period (commonly), without discounting.

Several items were needed to perform a BIA, namely:

- Dialysis coverage (%)
- Annual incidence rate (number of new patients)
- Annual prevalence rate (number of all patients)

A template to perform BIA was classified into a PD First Policy and a HD First Policy, shown in the table below.

Cohort 1 shows the current condition by calculating total prevalence rate, while Cohort 2 and beyond show the increasing number of dialysis cases (incidence rate) in the 2_{nd} year, and following years, respectively. BIA was performed by adjusting the number of current cases (Cohort 1) and costs related to the increased number of cases (Cohort 2, and so on).

Template Sample of BIA

Time Period (Year)	HD (Hemodialysis)								BIA
	Cohort 1/now	Cohort 2/new	Cohort 3/new	Cohort 4/new	Cohort 1/new	Cohort 2/new	Cohort 3/new	Cohort 4/new	
1									
2									
3									
4									
5									

For BIA, we changed the number of patients on the simulation sheet of both PD and HD in previous analysis, based on total prevalence and incidence rate within five years. This calculation was then inserted into the BIA template. Here, both PD and HD were analyzed, despite having a PD First Policy, or vice versa (HD First Policy). The result is shown in **Table II-9** below.

Table II-9. BIA Results for Dialysis

BIA within five years (in IDR)				
	PD First Policy		HD First Policy	
Year	Cost of PD	Cost of HD	Cost of PD	Cost of HD
1st	12.6 trillion			15.0 trillion
2nd	8.1 trillion	5.3 trillion	1.4 trillion	23.9 trillion
3rd	7.6 trillion	7.2 trillion	3.0 trillion	31.2 trillion
4th	7.5 trillion	8.6 trillion	4.5 trillion	37.5 trillion
5th	7.8 trillion	10.0 trillion	6.0 trillion	43.2 trillion
Total	43.5 trillion	31.0 trillion	14.8 trillion	150.9 trillion

Conclusions drawn from the BIA were as follows:

- HD and PD are complementary to each other, and one cannot be replaced by the other. Therefore, both therapies were analyzed, despite having only one as the first policy.

- A HD First Policy was predicted to have a IDR 91.2 trillion higher cost than a PD First Policy.

- In the PD First Policy, the highest cost was observed in the first year since it included the initial surgery, followed by lower costs in subsequent years. In contrast, the HD cost increased, due to the necessity of using dialysis machines and HD fluid, as there was an increasing number of patients.

Sample 2: BIA on pulmonary arterial hypertension (PAH) treatment using sildenafil

The economic evaluation result suggested that sildenafil would be a cost-effective treatment for PAH, with lower incremental cost-effectiveness ratio (ICER) between sildenafil and beraprost compared to the Indonesian GDP per capita threshold, i.e., sildenafil was considered to provide value for money, in functional class type II and III. BIA was performed using the same Microsoft Excel template that was used for economic evaluation. Several additional data were needed for the BIA, namely:

- a. Prevalence rate of PAH in Indonesia;
- b. Incidence rate of PAH in Indonesia;
- c. Total population of Indonesia;
- d. Proportion of PAH patients treated in functional class type II;
- e. Proportion of PAH patients treated in functional class type III

These data were required to determine the number of patients in the first and following years, over a five-year time span, as illustrated in **Table II-10**.

Table II-10. BIA for Sildenafil (in Millions of Rupiah)

Year	FC II		FC III	
	Beraprost	Sildenafil	Beraprost	Sildenafil
1	119.4	115.5	91.4	104.3
2	96.3	95.9	82.2	90.1
3	98.8	100.9	85.8	93.8
4	102.2	105.9	88.7	98.3
5	106.1	110.7	91.5	102.7
Total	**522.9**	**528.9**	**439.6**	**489.3**

The BIA results suggested that an additional IDR 55.7 million would be needed if BPJS substituted sildenafil for beraprost in the benefit package for PAH treatment.

Reporting BIA

The BIA results of an HTA are reported along with the economic evaluation results of the proposed health technology (CEA).

Decision-makers refer to the BIA results to decide on approval of the proposed benefit package, with consideration to equity, as well as social, ethical, and political issues, etc. Therefore, it is important to have a clear and rational BIA report so that decision-makers can easily understand it. It is highly recommended to mention all assumptions used in the report. If a proposed health intervention has limited data due to small numbers of cases, the assumptions employed have to be agreed upon by the relevant experts. The availability of registry data is considered to be very useful to produce credible BIA estimations.

The previously mentioned BIA sample on the PD First Policy for kidney failure treatment needs to be complemented by recommendations for

implemention, including a pilot study in certain areas to better plan supply-side readiness (such as physician training, distribution of PD fluid into remote areas, etc).

In sample 2, the recommendation to include sildenafil for PAH treatment into the JKN benefit package needs further steps, such as encouraging the pharmaceutical industry to register sildenafil at the POM Agency (Indonesian National Agency of Drug and Food Control).

Eventually, economic evaluation and BIA as a whole are delivered to the stakeholders according to HTA procedures that are regulated by the appropriate institutions in Indonesia. Subsequently, the stakeholders appraise the HTA results to make policy decisions.

References

1   Arnold RJG. Pharmacoeconomics, from theory to practice. CRC Press; 2010.

2   Andrew B, Sculpher M. An introduction to Markov Modeling for econo-mic evaluation. Pharmacoeconomics 24.4 (2006): 355-371.

3   BMJ. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ 2013;346:f1049 doi: 10.1136/bmj.f1049 (Published 25 March 2013)

4   Briggs A, Claxton K, Sculpher M. Decision modeling for health economic evaluation. Oxford University Press; 2006.

5   Drummond, Michael F., et al. Methods for the economic evaluation of health care programmes. Oxford university press, 2015.

6   Gray A, Clarke P, Wolstenholme J, Wordsworth S. Applied methods of cost-effectiveness analysis in health care. Oxford University Press, 2010.

7   Guidelines for health technology assessment in Thailand. Edisi ke-2. J Med Assoc Thailand; 2014: Suppl.5

8   Health care cost, quality, and outcomes: Ispor book of terms. International Society for Pharmacoeconomics and Outcome Research; 2003.

9   ISPOR. Budget impact analysis. 2012.

10   Keputusan Menteri Kesehatan Republik Indonesia Nomor 171/Menkes/ SK/IV/2014 tentang Komite Penilaian Teknologi Kesehatan.

11   Laporan hasil PTK: Efektivitas klinis dan evaluasi ekonomi terhadap terapi dialisis bagi pasien gagal ginjal terminal di Indonesia, Kementerian Kesehatan, 2016 (unpublished).

12   Laporan hasil PTK: Efektivitas dan Evaluasi Ekonomi Pemberian Sildenafil untuk Terapi Hipertensi Arteri pulmonal di Indonesia, Kementerian Kesehatan, 2016 (unpublished).

13   Matthew T. What is sensitivity analysis? Tersedia di: www.whatisseries. co.uk,. Hayward Medical Communication 2009.

14   Mauskopf JA, Sullivan SD, Annemans L, Caro J, Mullins CD, Nuijten M, et al. Principles of good practice for budget impact analysis: report of the ISPOR Task Force on Good Research Practices— Budget Impact Analy-sis. Value Health 2007; 10: 336-47.

15   Peraturan Presiden No.12 tahun 2013 tentang Jaminan Kesehatan

16 Rascati KL. Essentials of pharmacoeconomics. Lippincott Williams Wilkins; 2009.

17 Russel LJ, Gold MR, Siegel JE, Weinstein MC. Cost effectiveness in health and medicine. Oxford University Press;1996

18 Szende A, Oppe M, Devlin N, penyunting. EQ-5D Value sets, inventory, comparative review,and user guide. Euroqol Group Monographs Vol. 2 2007. ISBN: 978-1-4020-5510-2 (Print) 978-1-4020-5511-9 (Online).

19 Schwarzer R, Rochau U, Saverno K, Jahn B, Bornschein B, Muehlberger N, dkk. Systematic review of cost-effectiveness threshold in ten countries accross four continents. J Comp Eff Res. 2015 Sep;4:485-504.

20 Tan-Torres Edejer T, Baltussen R, Adam T, Hutubessy R, Acharya A, Evans DB, et al. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization; 2003.

Australia Indonesia Partnership
for Health Systems Strengthening
(AIPHSS)

THE ACCESS AND
DELIVERY PARTNERSHIP